

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGCGGTCTGGTAAGGATTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTGTGAAGTTGAAACAGAAAACCTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCTCAGTTCTTCCCTCAGCCCTTGTAATTGGACATCTGCTGCTTTCATATT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTACCTATATCAGTACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTGGGCAATGCTAAATATTGCGGCAGTTTATGCATTG
CTACCATTATGTTCGTTAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAA
TTAAACAAGGCTGGCCTTGTACTTGGAAACTGAGTTGTTAGGACTTCTATTGTGGCAAACCTT
CCAGAAAACAACCCCTTTGCTGCACATGTAAGTGGAGCTGTGCTTACCTTGGTATGGCCTCAT
TATATATGTTGTTCAAGACCACCTTCCTACCAAATGCAGCCCCAAATCCATGGCAAACAGTC
TTCTGGATCAGACTGTTGTTGTTATCTGGTGTGGAGTAAGTGCACCTAGCATGCTGACTTGCTC
ATCAGTTTGACAGTGGCAATTGGACTGATTAGAACAGAAACTCCATTGGAACCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAACGGTCTATGTCATTTCCTTCTT
GGTTTTCTGACTTACATTGGTGAATTTCAGAAAATTCTTACGGGTGGAAGCCAATTACA
TGGATTAACCCCTATGACACTGCACCTTGCCTATTAAACAATGAACGAACACGGCTACTTCCA
GAGATATTTGATGAAAGGATAAAAATTCTGTAATGATTGATTCTCAGGGATTGGGAAAGG
TTCACAGAAGTTGCTTATTCTCTGTGAAATTTCACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGGCCATTGATAGATTATTCTAAAGGATATCAT
CAAGAAGACTATTAAAAACACCTATGCCTATACCTTTTATCTCAGAAAATAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIFSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMILNIAAV
LCIATIYVRYKQVHALSPEENVIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQVFIRLLLVIWCGVSALSMLTCSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFFLTYIRDQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGCGGACGCGTGGGGAGAGCCGCAGTCGGCTGCAGCACCTGGAGAAGGCAGACC
GTGTGAGGGGCCGTGGCCCCAGCGTGTGGCCTCGGGAGTGGAAAGTGAGGCAGGAGCCTTC
CTTACACTTCGCCATGAGTTCTCATCGACTCCAGCATCATGATTACCTCCAGATACTATTTTG
GATTGGGTGGTTTCTTCATGCCAATTGTTAAAGACTATGAGATACTGTCAGTATGTTACAG
GTGATCTCTCCGTGACGTTGCATTCTGCACCAGTTGAGCTCATCATCTTGAAATCTTAGG
AGTATTGAATAGCAGCTCCGTTTTCACTGGAAATGAACCTGTGTAACTCTGCTGATCCTGG
TTTCATGGTGCCTTTTACATTGGTATTTATTGTGAGCAATATCCACTACTGCATAAACACGA
CTGCTTTTCCTGCTCTTATGGCTGACCTTATGTATTCTTCTGAAAATAGGAGATCCCTTCC
CATTCTCAGCCAAAACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGTTGGTGTGATTGGAG
TGACTCTCATGGCTCTTCTGGATTGGCTGTCAACTGCCATACACTTACATGTCTTACTTC
CTCAGGAATGTGACTGACACGGATTCTAGCCCTGGAACGGCAGTGTGCAAACCATGGATATGAT
CATAAAGCAAAAGAAAAGGATGGAATGGCACGGAGAACATGTCAGAAGGGAAAGTGCATAACA
AACCATCAGGTTCTGGGAATGATAAAAGTGTACCACTTCAGCATCAGGAATGAAAATCTTACT
CTTATTCAACAGGAAGTGGATGCTTGGAAAGAATTAAGCAGGCAGCTTTCTGAAACAGCTGATCT
ATATGCTACCAAGGAGAGAATAGAATACTCCAAACCTCAAGGGAAATATTTAATTTCTTGGTT
ACTTTCTCTATTACTGTGTTGGAAATTTCATGGCTACCATCAATATTGTTTGATCGAGTT
GGGAAACGGATCCTGTACAAGAGGCATTGAGATCACTGTGAATTATCTGGAATCCAATTGATGT
GAAGTTGGTCCAACACATTCCATTCTGGAAATAATCGTCACATCCATCAGAGGAT
TGCTGATCACTTACCAAGTCTTTATGCCATCTAGCAGTAAGTCTCCAATGTCATTGTCCTG
CTATTAGCACAGATAATGGCATGTACTTGTCTCTGTGCTGATCGAATGAGTATGCCCTT
AGAATACCGCACCATAACTGAAGTCCTGGAGAACTGCAGTTCAACTCTATCACCGTTGGTTG
ATGTGATCTCCTGGTCAGCGCTCTCTAGCATACTCTTCTCTATTGGCTCACAAACAGGCACCA
GAGAAGCAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTCAAATTAA
GATATAAGAGGGGGAAAATGAAACCAGGGCTGACATTATAACAAACAAAATGCTATGGTAGC
ATTTTCACCTCATAGCATACTCCTTCCCCGTAGGTGATACTATGACCATGAGTAGCATCAGCCAG
AACATGAGAGGGAGAACTAACTCAAGACAATACTCAGCAGAGGCATCCGTGAGGATATGAGGCTGG
TGTAGAGGGAGAGGGAGCCAAAGAAACTAAAGGTGAAAATACACTGGAACCTGGGGCAAGACATGT
CTATGGTAGCTGAGCCAACACGTAGGATTCCGTTTAAGGTTCACATGGAAAAGGTTAGCTT
CCTTGAGATTGACTCATTAAACAGAGACTGTAACAAAAAAAAAAAAAGGGCGGCCCG
ACTCTAGAGTCGACCTGCAGAAGCTGGCCGCATGGCCAACTTGTATTGCAAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQLFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
LNSSSRYFHWKMNLCVILLILVMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTMYFFWKLGD
FPILSPKHGILSIEQLISRVGVIGVTMALLSGFGAVNCPTYMSYFLRNVTDTDILALERRLLQ
TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWMGIKSVTTSASGSENLTLIQQEVDALELSRQ
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTDPVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLAQIMGY
FVSSVLLIRMSMLEYRTIITEVLGELQFNFYHRWFDV1FLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTAACATAGTTCC
AGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCATATCACCAAGTGGCCATCTGAGGT
GTTCCCTGGCTCTGAAGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT
TCCATCTGGACCACGAGGCTCTGGCCAAGGCTTTGCGTGAGAAGAGCTTCCATCCAGGT
GTCATGCAGAATTATGGGATCACCCCTGTGAGCAAAAAGGCGAACAGCAGCTGAATTTCACAG
AAGCTAAGGAGGCCTGTAGGCTGGACTAAGTTGCCGGCAAGGACCAAGTTGAAACAGCC
TTGAAAGCTAGTTGAAACCTGCACTATGGCTGGGGATGGAGATGGATTGTCATCTCTAG
GATTAGCCAAACCCAAAGTGTGGGAAAAATGGGGTGGGTGCTGATTGAAAGGTTCCAGTGA
GCCGACAGTTGCACTGCATTGCAACTCATCTGATACTTGACTAACCGTGCATTCCAGAA
ATTATCACCACCAAAGATCCCATTCAACACTCAAACAAACACAAGAATTATGTC
CAGTGACAGTACCTACTCGGTGGCATCCCTACTCTACAAATACCTGCCCCACTACTCTC
CTGCTCCAGCTCCACTTCTATTCCACGGAGAAAAATTGATTGTCACAGAAGTTTATG
GAAACTAGCACCCTGTCTACAGAAACTGAACCATTTGTTGAAAGAAAGCAGCATTCAAGAATGA
AGCTGCTGGGTTTGGAGGTGCCCCACGGCTGCTAGTGTGCTCTCCTTCTTGGTGTG
CAGCTGGCTTGGATTTCGCTATGTCAGGAAAGGTATGTGAAAGGCTTCCCTTTACAAACAAGAAT
CAGCAGAAGGAAATGATCGAAACAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCTAA
TGAGGAATCAAAGAAAATGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAACCTACCGTGC
GATGCCCTGGAAAGCTGAAGTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTCTT
CATGCTCTTACCCCTGCCAGCTGGGAAATCAAAGGCAAAGAACCAAAGAAGAAAGTCCA
CCCTGGTTCTAACTGGAATCAGCTCAGGACTGCCATTGACTATGGAGTGCACCAAAGAGAAT
GCCCTTCTCTTATTGTAACCTGTCTGGATCCTATCCTCACCTCAAAGCTCCCACGGCCT
TTCTAGCTGGCTATGCTTAATAATATCCACTGGGAGAAAGGAGTTTGCAAAGTGCAGGAC
CTAAACATCTCATCAGTATCCAGTGGTAAAAGGCCCTCTGGCTGTGAGGCTAGGTGGTTG
AAAGCCAAGGAGTCAGTGGACCAAGGCTTCTACTGATTCGCAGCTCAGACCCCTTCTCA
GCTCTGAAAGAGAAAACCGTATCCCACCTGACATGCTCTGAGCCGGTAAGAGCAAAGAAT
GCCAGAAAAGTTAGCCCCGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTGTAAA
GCTAAATAAGAAATAGAACAGGCTGAGGATACGACAGTACACTGTCAGCAGGGACTGAAAC
ACAGACAGGGTCAAAGTGTCTCTGAACACATTGAGTTGAATCACTGTTAGAACACACACA
CTTACTTTCTGGCTCTACCACTGCTGATATTCTAGGAAATATACTTTACAAGTAACA
AAATAAAAATCTTATAAAATTCTATTCTGAGTTACAGAAATGATTACTAAGGAAGATT
ACTCAGTAATTGTTAAAAGTAATAAAATTCAACAAACATTGCTGAATAGCTACTATATGTC
AAAGCTGTGCAAGGTATTACACTCTGTAATTGAATATTCTCCTCAAAAATTGCACATAGTAG
AACGCTATCTGGGAAGCTATTCTGAGTTGATATTCTAGCTATCTACTTCCAAACTAAT
TTTTATTCTGCTGAGACTAATCTTATTCTAATATGGCAACCATTATAACCTTAATT
TATTATTAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAGCACATTAAAAGTGC
ATTAACAAATGTATCACTAGCCCTCTTTCCAACAAGAAGGACTGAGAGATGCAGAAATT
TGTGACAAAAATTAAAGCATTAGAAAATT

FIGURE 6

MARCFSLVLLTSIWTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG
LSLAGKDQVETALKASFETCSYGVGDGFVVISRISPNCGKNGVGVLIWKPVRQFAAYCYN
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTPPAPASTSIPR
RKKLICVTEVFMETSTMSTETEPFVENKAASFNEAAGGGVPTALLVLALLFFGAAAGLGFCYVK
RYVKAFFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCCGCACCGCGGGCCGCCACCGCGCCGCTCCGCATCTGCACCCGAGCCCCGC
GGCCTCCCGCGGGAGCGAGCAGATCCAGTCGGCCCGAGCGCAACTCGGTCCAGTCGGGGCG
CGGCTCGGGCGCAGAGCGGAGATGCAGCGCTTGGGGCCACCCCTGCTGTGCCTGCTGGCG
CGGCGGTCCCCACGCCCGCGCCGCTCCGACGGCACCTCGCTCCAGTCAGCAAGCCGGCC
GCTCTCAGCTACCCGAGGAGGCCACCTCAATGAGATGTTCCGAGGGTTGAGGAACGTGAT
GGAGGACACGAGCACAATTGCGCAGCGGGTGAAGAGATGGAGGCAGAAGAAGCTGCTGTA
AAGCATCATCAGAAAGTGAACCTGCAAACCTACCTCCCAGCTATCACAAATGAGACCAACACAGAC
ACGAAGGTTGAAATAATACCATCCATGTGCACCGAGAAATTCAAGATAACCAACAACCAGAC
TGGACAAATGGTCTTTCAGAGACAGTTACATCTGTGGGAGACGAAGAAGGCAGAAGGGAGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCAGAGCATGTACTGCCAGTTGCCAGCTCCAGTAC
ACCTGCCAGCATGCCGGGCCAGAGGATGCTGCAACCGGGACAGTGAGTGCTGTGGAGACCA
GCTGTGTCTGGGTCACTGCACAAAATGCCACCAAGGGCAGCAATGGACCATCTGTGACA
ACCAGAGGGACTGCCAGGCCGGCTGTGCTGTGCCCTCAGAGAGGGCTGCTGTCCCTGTGTC
ACACCCCTGCCGTGGAGGGCGAGCTTGCATGACCCGCCAGCCGCTTCTGCCACCTCATCAC
CTGGGAGCTAGAGCCTGATGGAGCCTGGACCGATGCCCTGTGCCAGTGGCCTCCTGCCAGC
CCCACAGCCACAGCCTGGTGTATGTGCAAGCCGACCTCGTGGGAGGCCGTGACCAAGATGGG
GAGATCCTGCTGCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCCTGGGGAGCCTGCCGCTG
CCGCCGCTGCACTGCTGGAGGGGAAGAGATTAGATCTGGACCAGGCTGTTGGTAGATGTGCAA
TAGAAATAGCTAATTATTTCCCCAGGTGTGCTTAGGCGTGGGCTGACCAGGCTTCTCCTA
CATCTCTCCAGTAAGTTCCCTCTGGCTGACAGCATGAGGTGTTGCAATTGTCAGCT
CCCCCAGGCTGTTCTCCAGGCTTCACAGTCTGGTCTGGAGAGTCAGGCAGGGTTAAACTGCA
GGAGCAGTTGCCACCCCTGTCCAGATTATTGGCTGCTTGCCCTACAGTTGCCAGACAGCCG
TTTGTCTACATGGCTTGTATAATTGTTGAGGGAGGAGATGAAACAATGTGGAGTCTCCCTC
TGATTGGTTTGGGAAATGTGGAGAAGAGTGCCTGCTTGCAACATCAACCTGGAAAAATG
CAACAAATGAATTTCACGCAGTTCTTCATGGGATAGGTAAAGCTGTGCCCTCAGCTGTG
AGATGAAATGTTCTGTCACCTGCAATTACATGTGTTATTCACTCCAGCAGTGTGCTCAGCTCC
TACCTCTGTGCCAGGGCAGCATTTCATATCCAAGATCAATTCCCTCTCAGCACAGCCTGGG
AGGGGTCATTGTTCTCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTCATCTGGTGTGACTCTAAGCTCAGTGTCT
CTCCACTACCCCACACCAGCTGGGCCACCAAAAGTGTCTCCCCAAAAGGAAGGAGAATGGGAT
TTTCTTGAGGCATGCACATCTGAATTAAGGTCAAACATAATTCTCACATCCCTCTAAAGTAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGGCAGCCGCTCTTAATGAAGACAATGAT
ATTGACACTGTCCTCTTGCAAGTTGACTAGTAACATTGAAAGGTATATGACTGAGCGTAGCA
TACAGGTTAACCTGCAGAACAGTACTTAGTAATTGAGGGCGAGGATTATAAATGAAATTG
AAAATCACTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGTGTAAATATGCGACTGCGAACACTGAACCTACGCCACTCCACAAATGAT
TTTCAGGTGTCATGGACTGTTGCCACCATGATTCACTCCAGAGTTCTTAAAGTTAAAGTTGCA
CATGATTGTATAAGCATGTTCTTGAGTTAAATTATGTATAAACATAAGTTGCAATTAGAA
ATCAAGCATAAATCACTCAACTGCAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEEAAAKASSEVNLANLPPSYHNETNTDTKVNNTIHVHREIHKITNNQTGQMVFSE
TVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRMLCTRSECCGDQLCVWGHC
TKMATRGNSNGTICDNQRDCQPGLCQAFQRGLLFPVCTPLPVEGELCHDPASRLLDLITWELEPDG
ALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQELEDLE
RSLTEEMALGEPAAAAALLGGEEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGCGGACCGTGGGGCTGTGAGAAAGTCCAATAAACATCATGCAACCCAC
GGCCCACCTGTGAACCTCGTGCCTAGGGCTGATGTGCGTCTCAGGGCTACTCATCAAAG
GCCTAATCCAACGTTCTGTCTCAATCTGAAATCTATGGGGCTCTGGGCTTCTGGACCCCTT
AACTGGGTACTGGCCCTGGGCCATGGTCCCTGGAGCCTTGCCCTTACTGGCCTT
CCACAAGCCCCAGGACATCCCTACCTCCCTTAATCTCGCCTTCATCCGCACACTCCGTTACC
ACACTGGGTATTGGCATTGGAGCCCTACCTGACCCCTGTGCAGATAGCCCCGGTCATCTTG
GAGTATATTGACCACAGCTCAGAGGAGTGCAGAACCTGTAGCCCGCTGCATCATGTGCTGTTT
CAAGTGCCTCGGTCTGGTCTGGAAAAAATTATCAAGTCTAAACCGCAATGCATACATCATGA
TGCATCTACGGGAAGAATTCTGTCTCAGCCTTGCATGCTACTCATGCGAAC
ATTGTCAGGGTGGTGTCTGGACAAAGTCACAGACCTGCTGCTGTTCTTGGGAAGCTGCTGGT
GGTCGGAGGCGTGGGGCTCTGCTTCTTCTCCGGTCGATCCGGGCTGGTAAAG
ACTTTAACAGCCCCCACCTCAACTATTACTGGCTGCCATCATGACCTCATCCTGGGGCCTAT
GTCATGCCAGCGCTTCTCAGCTTCCGGATGTGTTGGACACGCTTCCCTGCTTCT
GGAAGACCTGGAGCGAACACGGCTCCCTGGACCGCCCTACTACATGTCCAAGAGCCTTCAA
AGATTCTGGCAAGAAGAACGAGGCGCCCCCGACAACAAGAAGAGGAAGTGACAGCTCCGG
CCCTGATCCAGGACTGCACCCCACCCCAACCGTCCAGCCATCCAACCTCACTTCCCTTACAGGT
CTCCATTGTGGTAAAAAAAGGTTTAGGCCAGGCGCCGTGGCTCACGCCGTAACTCCAACACT
TTGAGAGGCTGAGGCCGGGATCACCTGAGTCAGGAGTCAGGAGACCAGGCCATGGCAACATGGTG
AAACCTCCGTCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCACTCCA
GCTACTCGGGAGGCTGAGGCAGGAGAACGCTGAACCCGGAGGCAGAGGTTGCAGTGAGCCGA
GATCGGCCACTGCACCTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAACAAACAAA
AAGATTTATTAAAGATATTGTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLFWTL
NWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKF1KFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN
IVRVVVLDKVTDLLLFFGKLLVVGGVGVLFFFSGRIPLGKDFKSPHLNYYWLPIMTSILGAY
VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYYMSKSLLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

GCCCCCGCCGCCGGCGCCGGGCGCCCGAAGCCGGAGCCACCGCATGGGGCCTGCCTGGGAGCCTGCCTTCCTGGGTGAGGGAGGGGCC
TCCCTGCTCAGCTGCCGTCTGCCCTCTGCCGCTCTGCCCTGCATCCTGTGCAGCTGCTGCCCGC
CAGCGCAACTCCACCGTGAGCCGCCATCTCACGTTCTCCTCTTCTGGGGTGTGGGTCCA
TCATTATGCTGAGCCGGCGTGGAGAGTCAGCTCTACAAAGCTGCCCTGGGTGTGTGAGGAGGGGCC
GGGATCCCCACCGTCTGCAGGGCACATCGACTGTGGCTCCCTGCTTGGCTACCGCGCTGTCTACCG
CATGTGCTTCGCCACGGCGGCCCTCTTCTCTTCTTACCCCTGCTCATGCTCGCGTACCG
GCCGGGACCCCCGGCTGCCATCCAGAATGGGTTGGTTCTTAAGTCCGTATCCTGGTGGCCTC
ACCGTGGGTGCCCTACATCCCTGACGGCTCTCACCAACATCTGGTCTACTTCGGCGTGTGGG
CTCCTCCTCTCATCCTCATCCAGCTGGTGCTGCTCATCGACTTGGCACTCCTGGAACCGGGT
GGCTGGGAAGGCCGAGGAGTGCAGTCCCGTGCCCTGGTACCGCAGGCCCTCTTCTTCACTCTCCTC
TTCTACTTGCTGTCGATCGCGCCGTGGCGCTGATGTTCATGTACTACACTGAGCCCAGCGGCTGCC
CGAGGGCAAGGTCTCATCAGCCTAACCTCACCTCTGTGTCTGCGTGTCCATCGTGTCTGC
CCAAGGTCCAGGACGCCAGCCAACTCGGGTCTGCTGCAGGCCCTGGTACATCACCCTACACCAGT
TTTGTACCTGGTCAGCCATCCAGTATCCAGAACAGAAATGCAACCCCCATTGCAACCCAGCT
GGGCAACGAGACAGTTGTGGCAGGCCAGGGCTATGAGACCCAGTGGTGGGATGCCCGAGCATTG
TGGGCCTCATCATCTCCCTGTGCACCCCTCTCATCAGTCTGCGCTCTCAGACCAACCGGCAGGTG
AACAGCCTGATGCAGACCGAGGAGTGCCACCTATGCTAGACGCCACACAGCAGCAGCAGCAG
GGCAGCCTGTGAGGCCAGGGCTTGACAAACGAGCAGGACGGCGTCACCTACAGCTACTCCTCTCC
ACTTCTGCCCTGGTGTGGCTGCCACTGCACGTATGATGACGCTACCAACTGGTACAAGCCGGTGAG
ACCCGGAAAGATGATCAGCACGTGGACGCCGTGGGTGAAGATCTGTGCCAGCTGGCAGGGCTGC
CCTCTACCTGTGGACCCCTGGTAGGCCACTCCTCCTGCCAACCGCGACTCAGCTGAGGCAGCCTCA
CAGCCTGCCATCTGGTGCCCTGCCACCTGGTGCCCTCGGCTCGGTGACAGCCAACCTGCCCTC
CCCACACCAATGCCAGGTGAGCCCCACCCCTGCCAGCTCCAGGACCTGCCCTGAGGGGGC
CTTCTAGTCGTAGTGCCTCAGGGTCCGAGGAGCATCAGGCTCTGCAGAGCCCCATCCCCCGCCAC
ACCCACACGGTGGAGCTGCCCTTCCCTCCCTGGTGTGCCATACTCAGCATCTGGATGAA
AGGGCTCCCTGTCTCAGGCTCCAGGAGCAGGGCTGCTGGAGAGAGCAGGGAACTCCACACAG
TGGGGCATCCGGCACTGAAGCCCTGGTGTCCCTGGTCACGTCCCCCAGGGACCCCTGCCCTTCTG
GACTTCGTGCCCTACTGAGTCATCAAAGACTTTCTAAACAAAGCCAGTGCCTGTTAAAAAAA

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVESQL
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFTLIMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHAWNQRWLGKAE
ECDSRAWYAGLFFFFTLLFYLLSIAAVALMFYYTEPSGCHEGKVFISLNLTFCVCVSIAAVLPKV
QDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIIFLLCTLFISLRSSDHQRQNSLMQTEECPPMLDATQQQQQQVAACEGRAFDNEQDGVTYSY
SFFHFCLVLASLHVMMTLTNWYKPGTRKMISTWTAVWVKICASWAGLLLWTLVAPLLLNRD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGCCGGGCCAGGAACCACCCGTTAAGGTGTCTTCTCTTAGGGATGGTGA
GGTTGGAAAAAGACTCCTGTAACCCTCCTCCAGGGATGAACCACCTGCCAGAAGACATGGAGAACG
CTCTCACCGGGAGCCAGAGCTCCCATGCTTCCTGCGCAATATCCATTCCATCAACCCCACACAA
CTCATGGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTCTGTTGTTGTCACCTTGACCTCTTATTGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTCATAT
TTTGATATATTCTCTGGCAGTTTCGATTAAAGTGTAAACTTGCATATGCTGTGTCAG
ACTGCGCCATTGGTGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTACTAGCAAAG
TGATCCTTCGAAGCTTTCTCAAGGGGCTTGGCTATGTGTCGCCATCATTCATTCA
CTTGCCTGGATTGAGACGTTGGCTGGATTCAAAGTGTACCTCAAGAAGCAGAAGAAGAAAA
CAGACTCCTGATAGTCAGGATGCTTCAGAGAGGGCAGCACTTACCTGGTGGCTTCTGATG
GTCAGTTTATTCCCCCTCCTGAATCGAAGCAGGATCTGAAGAAGCTGAAGAAAACAGGACAGT
GAGAAACCACTTTAAGACTATGACTACTTTGTTAAATGTGAAAAACCCTCACAGAAAGTC
ATCGAGGCAAAAGAGGCCAGGCACTGGAGTCTCCCTGTCGACAGTAAAGTGTGAAATGGTAC
CACTGCTGGCTTATTGAACAGCTAATAAGATTATTGTAATACCTCACAAACGTTGTAC
CATATCCATGCACATTAGTGCCTGCCTGTTGCTGGTAAGGTAATGTCATGATTCACTCTCT
TCAGTGAGACTGAGCCTGATGTTAAACAAATAGGTGAAGAAAGTCTGTGCTGTATTCTAATC
AAAAGACTTAATATATTGAAGTAACACTTTTAGTAAGCAAGATACTTTTATTCAATTCA
AGAATGGAATTTTTGTTCATGTCTCAGATTATTGTATTCTTTAACACTCTACATT
TCCCTGTTTTAACTCATGCACATGTCCTTGTACAGTTAAAAGTGTAAATAAAACTG
ACATGTCAATGTTGGTAGTTTATTCTGTTGCATTATGTGTATGGCCTGAAGTGTGGA
CTTGCAAAAGGGAAAGAAAGGAATTGCGAATACATGTAAGGATGTCACCAGACATTGTATTATT
TTATCATGAAATCATGTTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTGAAATG
ACAAAATGACTAAACCATTCATGTTCCCTTGCAGCCAATTCAATTAAAATGAA
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQIMARIESYEGREKKGISDVRRTFCLFVTFDLLF
VTLLWIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFLDFKVLQPQEAEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTCGGGACCAGGACCCCCCTCGGGCCCGACCCGCCAGGAAAGACTGAGG
CCGGGGCCTGCCCCGGCGCTCTGGCGCCGCTCCGGGACAGAAGATGTGCTCC
GGTCCCCTGCTGCTGCTGGCGCTGCTGACTGGCCCTGGGGCTGGGGTGCAGGGCTGCCCAT
CCGGCTGCCAGTGAGCCAGCACAGACAGTCTCTGCACTGCCGCCAGGGGACCAAGGTGCC
CGAGACGTGCCACCCGACACGGTGGGCTGTACGTTGAGAACGGCATCACCATGTCAGCG
AGGCAGCTTGCGGGCTGCGGGGCTGCAGCTCCTGGACCTGTACAGAACAGATGCCAGCC
TGCCCAGGGGTCTCCAGCCACTGCCAACCTCAGCAACCTGGACCTGACGCCAACAGGCTG
CATGAAATCACCAATGAGACCTTCCGTGGCTGCGGCCGCTCGAGGCCCTAACGGCAAGAA
CCGCATCCGCCACATCCAGCCTGGTGCCTCGACAGCTGACCGCCTCTGGAGCTCAAGCTGC
AGGACAACGAGCTGCGGGACTGCCCCCGCTGCGCCTGCCGCCGCTGCTGCTGGACCTCAGC
ACAACAGCCTCTGGCCCTGGAGCCGGCATCCTGGACACTGCCAACCTGGAGGCGCTGCC
GGCTGGCTGGGGCTGCAGCAGTGGACAGGGCTCTCAGCCGCTGCGAACCTCCAGGACC
TGGATGTGTCGACAAACAGCTGGAGCGAGTGCACCTGTGATCCGAGGCCCTGGGGCTGACG
CGCCTGGGCTGGCCGGCAACACCCGCAATTGCCAGCTGCGGCCGAGGACCTGGGGCTGGC
TGCCCTGAGGAGCTGGATGTGAGCAACCTAACGCTGAGGCCCTGCCGGACCTCTGGGCC
TCTTCCCCGGCTGCGCTGCTGGCAGCTGCCGCAACCCCTCAACTGCGTGTGCCCTGAGC
TGGTTGGCCCTGGGTGCGAGAGGCCACGTACACTGGCAGGCCCTGAGGAGACGCCGTGCCA
CTTCCCGCCAAGAACGCTGGCCGGCTGCTCTGGAGCTTGACTACGCCGACTTGGCTGCCAG
CCACCACCAACAGCCACAGTGCCCCACACGCCAGGCCCCGGTGGTGCGGGAGGCCACAGCCTTGCT
TCTAGCTGGCTCCTACCTGGTTAGCCCCACAGGCCGGCCACTGAGGCCCTGGCCACCCCTGCC
CACTGCCCCACCGACTGTAGGGCTGTCCCCCAGCCCCAGGACTGCCCCACCGTCCACCTGCC
ATGGGGGACATGCCACCTGGGACACGGCACCCACTGGCGTGTGGCCGAAGGCTTCACG
GCCCTGTAACTGTGAGAGCCAGATGGGCGAGGGGACACGCCAGGACTACACCAGTCACGCCAG
GCCACACGGTCCCTGACCCCTGGCATCGAGCCGGTGAGCCCCACCTCCCTGCCGTGGGGTGC
AGCCTACCTCCAGGGGAGCTCGTGCAGCTAGGAGGCCCTCCGTCACCTATGCCAACCTATCG
GGCCCTGATAAGCGGCTGGTGAGCGACTGGCCACCTGGCTGTGCTGAGTACACGGTCACCCA
GCTGCCGCCAACGCCACTTACTCCGTCGTGTCATGCCCTGGGGCCGGGGTGCCGGAGG
GCGAGGAGGCCCTGGGGGAGGCCATACACCCCGCCACTCCAACCCACGCCCTGGGGTGC
CAGGCCGCCGAGGGCAACCTGCCCTCCATTGCCGCCGCCCTGGGCCGGTGCCTCCGCC
GCTGGCTGCCGTGGGGCAGCTACTGTGTCGGCGGGGGCGGGCCATGGCAGCAGCGGCTCAGG
ACAAAGGGCAGGTGGGGCAGGGCTGGGCCCTGGAACTGGAGGGAGTGAGGTCCCCTGGAG
CCAGGCCGCCGAAGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGCGGCTGAGGTGTGAGGTGCC
ACTCATGGCTTCCAGGGCTGGCTCCAGTCACCCCTCCACGCAAGCCCTACATTAAGCCA
GAGAGAGACAGGGCAGCTGGGGCAGGCCCTGAGCCAGTGAGATGCCAGCCCCCTCTGCTGCC
ACACCACTGTAAGTCTCAGTCCAACCTGGGGATGTGTCAGACAGGGCTGTGACCACAGCT
GGGCCCTGTTCCCTGACCTGGCTCCATCTGAGATGCTGAGGCCAGCTGACGCC
CTAACGTCCCCCAGAACGAGTGCTATGAGGACAGTGTCCCGCCCTGCCCTCCGCAAGTGCAGT
CCTGGGCACGGCGGGCCCTGCCATGTGCTGTAACGCATGCTGGGCTCTGCTGGCTCT
TCCAGGGGGACCCCTGGGGGCCAGTGAGGAAGCTCCCGAAAGAGCAGAGGGAGACGCC
GGCTGTGACTCTAGTCTGGCCCCAGGAAGCGAAGGAACAAAAGAAAACTGGAAAGGAAGATGC
TTAGGAACATGTTTGCTTTTAAATATATATTATAAGAGATCCTTCCCATTTATTCT
GGGAAGATGTTTCAAACTCAGAGACAAGGACTTGGTTTGTAAGACAACCGATGATGATGAA
GGCCTTGTTAAGAAAAAATAAAAGATGAAAGTGTGAA

FIGURE 16

MCSRVPLLLPLLLLLALGPGVQGCPSCQCSCPQTVFCTARQGTTVPRDVPPDTVGLYVFENGIT
MLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRLERLY
LGKNRIRHIQPGAFDTLDRLLELKLQDNELRALPPLRIPRLLLDLSHNSLLALEPGIILDNTANVE
ALRLAGLGLQQLDEGLFSRLRNLIHDLDVSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL
AGLAALQELDVSNLSQLALPGDLSGLFPRLLAARNPNCVCPLSWFGPWVRESHVTLASPEE
TRCHFPPKNAQRLLLELDYADFGCPATTTTATVPTTRPVVREPTALSSSLAPTWLSPTAPATEAP
SPPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGTRHHACLCPEGFTGLYCESQMGGTRPSPTP
VTPRPPRSLTGLIEPVSPSLRVGLQRYLQGSSVQLRSLRTRYRNLSGPDKRLVTLRLPASLAEY
TVTQLRPNATYSVCVMPLGPGRVPEGEEACGEAHTPPAVHSNHAPVTQAREGNLPLLIAPALAAV
LLAALAAVGAAAYCVRGRAMAAAQDKGQVGPGAGPLELEGVKVPLEPGPKATEGGEALPSGSE
CEVPLMGFPGPGLQSPLHAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

GCAGCGCGAGGCAGCGGTGGCTGAGTCGTGGTGCAGAGGCGAAGGCGACAGCTCATGCG
GGTCCGGATAGGCAGCAGGATGAATCCTAGATTCAAGACTACTTGCACATCAGATGAGTCAGTA
ATGAAGAAGGCAGCAGGATGAATCCTAGATTCAAGACTACTTGCACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTCCTGATTCAAAGAAATCTGA
ATTAGAACCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGGAAAGTGTACAG
AAGATATCAGCTTCTAGAGTCAGGAAATCCAGAAAACAAGGACTATGAAGAGCCAAGAAAAGTA
CGGAAACCCAGCTTGACGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTCCCTTCT
TTCCCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGAAAGATGGCAGACTGTGGTGT
CTACAACCTATGACTACAAAGCAGATGAAAAGTGGGCTTTGTGAAACTGAAGAAGAGGCTGCT
AAGAGACGGCAGATCAGGAAGCAGAAAATGATGTATCAAACCTGAATGAAAATCCTTAATGAAAG
CAATAAGAAAAGCCAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTATGCTCTTTATTGGTGTGAAACTTGCCACAGAATATCCAG
GCAGCGAGAGAGATTTGAGAAGCTGACTGAGGAAGGCTCTCCAAAGGACAGACTGCTCTGG
CTTCTGTATGCCTCTGGACTTGGTGTAAATTCAAGTCAGGCAAAGGCTCTGTATATTACAT
TTGGAGCTCTGGGGCAATCTAAAGCCCACATGGTTTGGTAAGTAGACTTTAGTGGAAAGGCT
AATAATATTAACATCAGAAGAATTGGTTTATAGCGGCCACAACCTTTCAGCTTCATGATC
CAGATTGCTGTATTAAGACCAAATTCAGTTGAACCTCCTCAAATTCTGTTAATGGATAT
AACACATGGAATCTACATGAAATGAAAGTGGTGGAGTCCACAATTCTTTAAATTTCTGAGTTG
TTGGCTGATTGCCCTAAAAGAGAGATCTGATAAAATGGCTCTTTAAATTTCTGAGTTG
GAATTGTCAGAATCATTTCACATTAGATTATCATAATTAAAATTTCTTAGTTCA
AAATTGTAATGGCTATAGAAAACAACATGAAATATTATACAATATTGCAACAATGC
CCTAAGAATTGTTAAATTGAGTTATTGTCAGAATGACTCCAGAGAGCTCTACTTCTG
TTTTTACTTTCATGATTGGCTGTCTCCATTATTCTGGTCAATTGTTACTTTCTTGCTAATTGG
GCCTGCTTCCAGTAGTCTCATTTCCATTGGTCAATTGTTACTTTCTTGCTAATTGG
AAGATTAACTCATTAAATAAAATTATGCTAAGATTAAGATTAAGATTAAGATTAAGATTAAG
AAA

FIGURE 18

MRVRIGLTLLLCAVLLSLASASSDEEGSQDESLDSTTTSDESVKDHTAGRVVAGQIFLDSEEESEL
ESSIQEEEDSLKSQEGERVTEDISFLESPPNPKDYEEPKKVRKPALTAIEGTAAHGPCHFPFLFLDK
EYDECTS DGDREDGRWCATTYDYKADEKWGFCETEEEAKRRQM QEAEMMYQTGMKILNGSNKKSQKR
EAYRYLQKAASMNHTKALERVSYALLFGDYLQPNIQAAREMFEKLTEEGSPKGQTALGFLYASGLGVN
SSQAKALVYYTFGALGGNLIAHMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCAGATTTAACCCATTCTGCAGTGGAAATTGACTGAACAGAAGGACACCATTCTT
GTATTATAACAAGAAAGGAGTGTACCTATCACACACAGGGGGAAAAATGCTTTGGGTGCTAGG
CCTCCTAATCCTCTGGTTCTGTGGACTCGTAAAGGAAAACTAAAGATTGAAGACATCACTG
ATAAGTACATTTTATCACTGGATGTGACTCGGCTTGGAAACTTGGCAGCCAGAAACTTTGAT
AAAAAGGGATTCATGTAATCGCTGCCTGCTGACTGAATCAGGATCAACAGCTTAAAGGCAGA
AACCTCAGAGAGACTCGTACTGTGCTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG
CCCAGTGGGTGAAGAACCAAGTGGGGAGAAAGGTCTCTGGGTCTGATCAATAATGGCTGGTGT
CCCGCGTGCTGGCTCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA
CCTGTTGGACTCATCAGTGTGACACTAAATGCTCCTTGGTCAAGAAAGCTCAAGGGAGAG
TTATTAATGTCCTCAGTGTGAGGTCGCCTGCAATCGTGGAGGGGCTATACTCCATCCAAA
TATGCAGTGGAAAGGTTCAATGACAGCTTAAGACGGACATGAAAGCTTGGTGCACGTCTC
ATGCATTGAACCAGGATTGTTCAAAACAAACTGGCAGATCCAGTAAAGGTATTGAAAAAAAC
TCGCCATTGGAGCAGCTGTCTCCAGACATCAAACAATATGGAGAGGTTACATTGAAAAA
AGTCTAGACAAACTGAAAGGCAATAATCCTATGTGAACATGGACCTCTCCGGTGGTAGAGTG
CATGGACCACGCTCAACAAAGCTCTTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAA
TTTCTGGATACCCTGTCTCACATGCCAGCAGCTTGAAGACTTTTATTGTTGAACAGAAA
GCAGAGCTGGCTAATCCCAAGGCAGTGTGACTCAGCTAACCCAAATGTCTCCAGGCTATGA
AATTGGCGATTCAAGAACACATCTCCTTCAACCCATTCTTATCTGCTCAACCTGGACT
CATTTAGATCGTGTCTTGGATTGCAAAAGGGAGTCCCCCACCATCGCTGGTGGTATCCAGGGT
CCCTGCTCAAGTTTCTTGGAAAGGAGGGCTGGATGGTACATCACATAGGCAAGTCCGTCCCT
GTATTTAGGCTTGGCTGCTGTGTGATGTAAGGGAATTGAAAGGACTTGCCATTCAAATG
TCTTACCGTGGCTGCCCATGCTTATGGTCCCCAGCATTACAGTAACTGTGAATGTTAAGT
ATCATCTTATCTAAATTAAAAGATAGTCAACCCAAAAAAAAAAAA
AAAAAAAAAAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGKLKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTEG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDWLTLEDY
REPIEVNLFGlisVTLNMLPLVKKAQGRVINSSVGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
AFGVHVSCIEPGLFKTNLADPVKVIKEKKLAIWEQLSPDIKQQYGEGYIEKSLDKLKGKNSYVNMD
LSPVVECMDHALTSLFPKTHYAAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGGCGCGGTAGCATGGAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTGTGCTCGCG
CACTCGCTTCCAGCACCTAACACGGACTCGGACACCGAAGGTTTCTTCTGGGAAGTAAAA
GGTGAAGCCAAGAACAGCATTACTGATTCCAAATGGATGATGTTGAAGTTGTTATACAATTGA
CATTCAAAATATATTCCATGCTATCAGTTTAGCTTATAATTCTTCAGGCGAAGTAATG
AGCAAGCACTGAAGAAAATATTCAAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATTCCGT
CGTCATTCAAGATCAGATCATGACGTTAGAGAGAGGCTGCTTCACAAAAACTTGCAAGGAGCATT
TTCAAACCAAGACCTGTTCTGCTATTAACACCAAGTATAATAACAGAAAGCTGCTCTACTC
ATCGACTGGAACATCCCTATATAAACCTCAAAAAGGACTTTACAGGGTACCTTAGTGGTT
GCCAATCTGGGCATGTCGAACAACGGTTATAAAACTGTATCAGGTTCTGTATGTCCACTGG
TTTAGCCGAGCAGTACAAACACAGCTCTAAATTGGAGAAGATGGATCCTTAAAGGAGG
TACATAAGATAAAATGAAATGTATGCTTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAGTG
GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAACAGATTAAACGAGAAATTGA
GAAAAGGAGAGGAGCACAGATTCAAGCAGCAAGAGAGAAGAACATCAGGAGA
ACATTTCCTTGTCAAGCATTACGGACCTTTTCCAAATTCTGAATTCTTCATTCAATGTGTT
ATGTCTTAAAAAAATAGACATGTTCTAAAAGTAGCTGTAACTACAACCACATCTGATGTAGT
AGACAATCTGACCTTAATGGTAGAACACACTGACATTCTGAAGCTAGTCCAGCTAGTACACCAC
AAATCATTAAGCATAAAGCCTAGACTTAGATGACAGATGGCAATTCAAGAGATCTGGTTGTTA
GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAACGATCCAAAAT
GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTGGTGAATATTACGGTCTC
CTACATTTGATCTTTAACCTTACAAGGAGATTTTATTTGGCTGATGGTAAAGCCAAAC
ATTCTATTGTTTACTATGTTGAGCTACTGCACTAAGTTCATTGTTTACTATGTTCA
TGTTGCACTAATACACAGATACTCTTAGTCATTACAAAGTACTTTCAACATCA
GATGCTTTATTCACAAACCTTTTCACTTCAACTAAGTTGTTGAGGGGAAGGCTTACACAG
ACACATTCTTGAATTGGAAAAGTGAGACCAAGGACAGTCAGTGGCTCACACCTGTAATCCCAGCACT
TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGCAACGTATT
GAGACCATGTCTATTAAAAAATGGAAAAGCAAGAATAGCCTTATTTCAAAATATGGAAA
GAAATTATGAAAATTATCTGAGTCATTAATTCCTTAAGTGATACTTTAGAAGTA
CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAATTGCAAAACATCATCT
AAAATTAAAAAAAAAAAAAAAAAAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTSDTEGFLLGEVKGEAKNSITDSQMDDVEVVYTIDIQKYI
PCYQLFSFYNSSGEVNEQALKKILSNVKKNVGWYKFRRHSDQIMTFRERLLHKNLQEHSNQDL
VFLLLTPSIITESCSTHRLEHSILYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSRAV
QTHSSKFFEDGSLKEVHKINEMYASLQEELKSICKVEDSEQAVDKLVKDVNRLKREIEKRGA
QIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCNYNHHLDVVDNLTL
MVEHTDIPEASPASTPQIIKHKA
LDLDDRWFQFKRSRLLDTQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGGCGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGACCAGCGCAGGGCAGCCAA
GCAGCGCGCAGCGAACGCCCGCCGCCACACCCCTCGCGTCCCGCGCCCTGCCACCCCTCCCTCCCTCCCC
GCGTCCCCCGCCTCGCCGCCAGTCAGCTTGCCGGGTCGCTGCCCGCGAAACCCCGAGGTACCAAGCCCGCCCTCT
GCTTCCCTGGGCCGCGCCGCCCTCACGCCCTCCTCTCCCTGGCCCGCCTGGCACCGGGACCGTTGCCGA
CGCGAGGCCAGCTACTTTGCCCGCTCTCCCGCCCTGCTCGCTCTCCACCCATCCAACTCCAACCTCCCTTCTCCC
TCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCGCTGCGTAGGCCGCTTCCCGTCCGGTCCAAA
GGTGGGAACCGTCCGCCCGGCCGACCATGGCACGGTCGGCTGCCCGCCTCTCTGCACCCGGCAGTGCTC
AGCGCCGCCGCTGGCTGCCGAGCTCAAGTCGAAAAGTGTGCTGGAAACTCGCACCGTCTTACGTGTC
AACAAAGAACGATGCCCTCCACGAGATCACGGTATCATTGAAGATCTGCCCCAGGGTCTACCTGCTGCTCT
CAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGATGATTCAAAGTGTGGTCAGCGAACAGTGAATCATTG
CAAGCTGCTTGCTTCACTGTTACAAGAAGTGTGAAATTCTCAAAGAAACTACTGAAAATGAGAATCCCTG
AATGATATGTTGTGAAGACATATGGCATTATACATGCAAATTCTGAGCTATTAAAGATCTTCTGAGAGTTG
AAACGTTACTACGTGGTGGAAATGTGAACCTGGAAGAAATGCTAAATGACTTCTGGCTCGCTCCTGGAGCGATG
TTCCGCCCTGGTA
CTCCAGTACCAAGTGTGAGCTTACAGATGAGTATCTGGAATGTGAGCAAGTATA
CCGGAGCAGCTGAAG
CCCTCGGAGATGCTCGCAAATTGAAGCTCCAGGTTACTCGTGTGTTGTAGCAGCCGTACTTCTGCTCAAGGC
TTAGCGGTTGCCGGAGATGCTGAGCAAGGTCTCGTGTGAAACCCAGGCCAGTGTACCCATGCCCTGTTGAAG
ATGATCTACTGCTCCACTGCCGGGTCTCGTGA
CTGACTGTGAGCCATGTTACA
ACTACTGCTCAAACATCATGAGAGGC
TGTTGCCCAACCAAGGGATCTCGATTGAAATGGAACAAATTCTAGATGCTATGCTGATGTTGGCAGAGAGGCTA
GAGGGTCCCTTCAACATTGAATCGGTATGGATCCCATCGATGTGAAGATTCTGATGCTATTGAAACATGCA
GAGGAT
AATAGTGTCAAGTGTCTCAGAAGGTTTCCAGGGATGTGGACCCCCAAGCCCTCCAGCTGGACGAATTCTCGT
TCCATCTGAAAGTGCCTCAGTGCTCGCTCAGACCCACATCCCCGAGGAACGCCAACACAGCAGCTGGCACT
AGTTGGACCGACTGGTACTGATGTCAGAGGAAACTGAAACAGGCCAAGAAATTCTGGCTCCCTCCGAGCAAC
GTTTGCAACGATGAGAGGATGGCTCGAGGAACGGCAATGAGGTGACTGTGGAAAGGCAAAAGCAGGTAC
CTGTTGCAGTGACAGGAATGGATTAGCCAACCAGGGCAACACCCAGAGGTCCAGGGTACACCAGCAAAC
ATACTGATCCTCGTCAAATCATGGCTTCTCGAGTGACCAAGATGAAAGATGCATACAATGGAACGACGTG
GACTCTTGATATCAGTGATGAAAGTGTGAGAAGGAAGTGGAAAGTGGCTGTGAGTATCAGCAGTGCCTTCAGAG
TTTGACTACAATGCCACTGACCATGCTGGAAAGAGTCCAATGAGAAAGCCGACAGTGCTGGTCCGTCTGGGCA
CAGGCCCTACCTCTCACTCTCTCATCTGTTCTGGTTATGCAAGAGAGTGGAGATAATTCTCAAACCTGAG
AAAAAGTGTGATCAAAAGTTAAAGGCACCAGTTACACTTTTACCATCCTAGTGACTTTGCTTTAAATGAA
TGGACAACAATGTACAGTTTACTATGTGGCACTGGTTAAGAAGTGTGACTTTGTTCTCATTGAGTTGGG
AGGAAAAGGGACTGTGCAATTGAGTTGGTCTGCTCCCCAACCATGTTAACAGTGGCTAACAGTGTAGGTACAGAA
CTATAGTTAGTTGTGCAATTGAGTTTATCACTCTATTATGTTGTTATGTTTTCTCATTGCTTGTGGGTT
TTTTTCTCAACTGTGATCTGCCCTGTTCTTACAAGCAAACAGGGTCCCTCTGGCACGTAA
CATGTACGTATT
TCTGAAATATTAAATAGCTGTACAGAACAGGTTTATTATCATGTTATTAAAGAAAAGCCAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAAELKSKCSEVRRILYVSKGFNKNDAPLHEINGDHLKICPQGST
CCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDFVKTYGH
LYMQNSELFKDLFVELKRYYVGPNVLEEMLNDFWARLLERMFRLVNSQYHFTDEYLECVSKYTE
QLKPFGDVPRLKLQVTRAFVAARTFAQGLAVAGDVSKVSVNPTAQCTHALLKMIYCSHCRLG
VTVKPCNYCSNIMRGCLANQGDLDFEWNNFIDAMLMAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKVFOQCGPPKPLPAGRISRSISESAFSARFRPHHPEERTTAAGTSLDRLVTDVK
EKLQAKKFWSLPSNCNDERMAAGNGNECDCWNGKGKSRYLFAVTGNGLANQGNPENVQVDT
KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGCAACT
TACAGCTGCACCGACAGTTGCGGATGAAAGTTCTAATCTCTTCCCTCCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCACAGGGACCGAGGC
CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAAGATTGGTTCTGAG
AGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAACAGCAGTGCCCCTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAGGCACACAGAAAGCCAACAAGCATTCCAGA
GCCTGCCAGCAATTCTCAAACAATGTCAGCTAAGAAGCTTGCTCTGCCTTTGTAGGAGCTTG
AGCGCCCACCTTCCAATTAAACATTCTCAGCCAAGAACAGTGAACACACCTACCAAGACACTC
TTCTTCTCCCACCTCACTCTCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAGCA
TGTTTTCAAGATCATTGTTGCTCTCTAGTGTCTTCTCGTCAGTCTTAGCCT
GTGCCCTCCCTAACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCC
AGCTAGTGTCAATTAAACCTAAATGCAATCAGGAAAGTAGCAAACAGAACGTCAATAATATT
AAATGTCAAAAAAAAAAAAAAAA

FIGURE 26

MKVLISLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKKQCPDCDFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTCCCCAGCGGAAGCACAGCTCAG
AGCTGGTCTGCCATGGACATCCTGGTCCCACCTGCAGCTGCTGGTCTGCTTCTTACCTGCC
CCTGCACCTCATGGCTCTGCTGGCTGGCAGCCCCGTGCAAAGACTACTTCCCTACCTGA
TGGCGTGCTGACTCCAAAGAGCAACCGCAAGATGGAGAGCAAGAACGGAGCTTCAGCAG
ATAAAGGGGCTTACAGGAGCCTCCGGAAAGTGGCCCTACTGGAGCTGGCTGGGAACCGGAGC
CAACTTCAGTTCTACCCACCGGGCTGCAGGGTACCTGCCTAGACCCAAATCCCCACTTGAGA
AGTTCTGACAAAGAGCATGGCTGAGAACAGGCACCTCAAATATGAGCGGTTGGTGGCTCCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGTG
CTCTGTGCAGAGCCAAGGAAGGTCTGCAGGAGGTCCGGAGACTACTGAGACCGGGAGGTGTG
TCTTTCTGGAGCATGTGGCAGAACATATGGAAGCTGGCCTTCATGTGGCAGCAAGTTTC
GAGCCCACCTGAAACACATTGGGATGGCTGCCTCACCAGAGAGACCTGGAAGGATCTGA
GAACGCCAGTTCTCCGAAATCCAATGGAACGACAGCCCCCTCCCTGAAGTGGCTACCTGTTG
GGCCCCACATCATGGGAAAGGCTGCAAACAATCTTCCAAGCTCCAAGGCACTCATTGCTCC
TTCCCCAGCCTCCAATTAGAACAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACTA
GCAGAATGAGAGAAGACATTCTGACAGTACCTACTAGTCCCTCTCCCAACCTCTGCCAGGGC
AATCTCTAACTCAATCCGCCTCGACAGTAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
AAACACTAGGACCCCTGTTGTATCCTCAACTGCAAGTTCTGGACTAGTCTCCCAACGTTGCCTC
CCAATGTTGTCCCTTCCTCGTCCATGGTAAAGCTCTCGCTTCTCTGAGGCTACAC
CCATGCGTCTCTAGGAACACTGGTACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGAC
CCTCTCTCCCCACTACCACCTCTTCTGAGCTGGGGCACCAGGGAGAATCAGAGATGCTGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTGTTCTCAAATTTTTAATAAATAGACGAA
ACCACG

FIGURE 28

MDILVPLLQLLVLLLTLPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSQIKGL
TGASGKVALLELGCGTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDVVVCTLVLCSVQSPRKVLQEVRRLRPGGVLFFWEHVAEPYGSWAFMWQQVFEPTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPGPHIMGKAVKQSFPSSKALICSFPSL
QLEQATHQPIYLPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATTTGCCTATCCACCTCCCCAAGCCCCTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTAAAGGCTCATGCTTGGAGTGGGACTGGTCGGTGCCAGAAAGTCTCTCTG
CCACTGACGCCCCCATCAGGGATTGGGCCTTCTTTCCCCCTCTTCTGTCTCCTGCCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGAAGGGAGAAAGTGGGGATGGCTA
AGAAAGCTGGAGATAGGAAACAGAACAGAGGGTAGTGGGTGGCTAGGGGGCTGCCTATTAAA
GTGGTTGTTATGATTCTTATACTAATTACAAAGATATTAAAGCCCTGTTCATTAAGAAATT
GTTCCCTCCCCGTGTTCAATGTTGTAAAGATTGTTCTGTGTAAATATGTCTTATAATAAAC
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MLLLTL^LLL^LLL^LLL^LKGS^CLEWGLVGAQKVSSATDAP^IRDW^AF^FPPSFLCLLPHRPAMTC^SQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTCGAATTCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAGTT
CCTCCAAGCAAGTCATTCCCTATTAACCGATGTGTCCTCAAACACCTGAGTGCTACTCCCT
ATTTGCATCTGTTTGTATAATGATGTTGACACCCCTCCACCGAATTCTAAGTGGAAATCATGTCGG
GAAGAGATAACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTGGCCATGATGTTACC
TTCAGATTCATCACCAACCCCTCTGGTCACATTTCATGGTTATTTGGGATTGTTGTT
TGTCTGCGGTGTTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA
CAGAAAGGGAAAATATGAAGTGGTGTGGGTTGCTATCGTATCCACAGGCATCACGGCAGTG
CTGCTCGTCTGATTTGGTCTCAGAAAAGAGAATAAAATTGACAGTTGAGCTTTCCAAATCAC
AAATAAAGCCATCAGCAGTGCTCCCTCCTGCTGTTCCAGCCACTGTGGACATTGCCATCCTCA
TTTCTCTGGTCCCTCTGGGTGGCTGTGCTGAGCCTGGAACTGCAGGAGCTGCCAGGTT
ATGGAAGGCAGGCAAGTGGAAATAAGCCCCTTCGGGCATTGGTACATGTGGTGTACCATTT
AATTGGCCTCATCTGGACTAGTGAATTCATCCTGCGTGCAGCAAATGACTATAGCTGGGCAG
TGGTTACTTGTATTCAACAGAAGTAAAATGATCCTCCTGATCATCCCATCCTTCGTCTCTC
TCCATTCTCTTCTTCAACATCAAGGAACCGTTGTGAAAGGGTCAATTAAATCTCTGTGGTGT
GATTCCGAGAACATTGTCACTGTACATGCAAAACGCAGTGAAAGAACAGCAGCATGGTGCATTGT
CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTCTGGTGTCTGACAAATACCTGCTCCATCTC
AACAGAACATGCATATACTACAACTGCTATTAAATGGGACAGATTCTGTACATCAGCAAAAGATGC
ATTCAAATCTGTCCAAGAACTCAAGTCACTTACATCTATTAACTGCTTGGAGACTTCATAA
TTTTCTAGGAAAGGGTTAGTGGTGTGTTCACTGTTGGAGGACTCATGGCTTTAACTAC
AATCGGGCATTCCAGGTGTGGCAGTCCCTCTGTTATTGGTAGCTTTGCCTACTTAGTAGC
CCATAGTTTATCTGTGTTGAAACTGTGCTGGATGCACTTTCTGTGTTGCTGTTGATC
TGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTATGGATCAAGAATTCTGAGTTGTA
AAAAGGAGCAACAAATTAAACATGCAAGGGCACAGCAGGACAAGCAGCTCATTAAGGAATGAGGA
GGGAACAGAACACTCCAGGCCATTGTGAGATAGAACCCATTAGGTATCTGTACCTGGAAAACATT
TCCTTCTAAGAGCCATTACAGAATAGAAGATGAGACCAGTAGAGAAAAGTTAGTGAATTTTTT
TTAAAAGACCTAATAAACCCATTCTTCCCTAAAA

FIGURE 32

MSGRTDTIILGLCILALALS LAMMFTFR FITTLLVHIFISLVLG LLLF VCGV LWL YYD YTN DLSIE
LDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTV ELFQITNKAIS SAPFLLFQPLWTFA
ILIFFWVLWVAVLLSLGTAGAAQVM EGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQQMTIA
GAVVTCYFNRSKNDPPDHPI LSSLSILFFYHQGTVVKG SFLISVVRIPRIIVMYMQNALKEQQHG
ALSRYLFRCYC CFWCLDKYLLHLNQNAYTTAINGTDFCTS AKA DFKILSKNSHFTSINC FGD
FIIFLGKV LVVCFTVFGGLMAFNYNRAFQVWAVP LLLVAFFAYLVAHSFLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTCGATTAGCTCCTGAGAAGAAGAGAAAAGGTCTGGACCTCCCTGTTCTCCTAGA
ATAATTGTATGGGATTGTGATGCAGGAAGCCTAACGGAAAAAGAATATTCAATTCTGTGTGGT
GAAAATTGGAAAAAAATTGCCTCTTCAAACAAGGGTGTCAATTCTGATAATTTATGAGGAC
TGTTGTTCACTATGAAGGCATCTGTATTGAAATGTCCTGTTGCTGGTACTGGAGTAC
ATTCAAACAAAGAACGGAAAGAAGATAAAAGGCCAAGTCACTGTGCCTCAGATCAACTGC
GATGTCAGAACCGGAAAGATCATCGATCCTGAGTCATTGTGAAATGTCAGCAGGGATGCCAAGA
CCCCAAATACCATGTTATGGCACTGACGTGTATGCATCCTACTCCAGTGTGTGGCGCTGCCG
TACACAGTGGTGTGTTGATAATTCAAGGAGGAAAATACTTGTCGAGGTTGCTGGACAGTCT
GGTTACAAGGGAGTTATTCCAACGGTGTCCAACTCGTTACCCCTACACAGATGGAGAGAATCCTT
TATCGTCTAGAAAGTAAACCCAAAAGGTTGTAACCTACCGCTTACATACTCATCAT
CGAAAAGTCCAGCTGCCAAGCAGGTGAGCACACAAAAGCCTATCAGAGGCCACCTATTCCAGGG
ACAACGTGACAGCCGGTCACTCTGTCAGTCAGCTCTGGCTGTCACTGTGCTGGCCACCCCCAC
CACCTTGCCAAGGCCATCCCCCTCTGTCGTTCTACCCAGCATCCCCAGACCACAATCAGTGG
GCCACAGGAGCCAGGAGATGGATCTGTCGTTCACTGCCACCTACAAAGCAGCCAAAACAGGCC
AGAGCTGATCCAGGTATCCAAGGCAAGATCCTTCAGGAGCTGCCTCCAGAAACCTGTTGGAGC
GGATGTCAGCTGGACTGTTCCAAAAGAAGAATTGAGCACACAGTCTTGGAGGCCAGTATCCC
TGGGAGATCCAAAAGTCAAAGCAGCTCTGCTGATGTTGCCAAGCTTGTGACATTGGCAAA
CGGCGATTCCGAATCCAGAACAGCAGCTCTGCTGATGGAGACAACCTGCTACTCACTTAACCTCAAGACAC
ACACGAATTCTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGGACTTCTAAT
GTAGGTCGGGCCATCTCTTGACCAAGAACATTCTTCCAAGGCAATGGAAACAGAACAGCGG
GGCTCCAAATGTGGTGGTGTGATGGTGGATGGCTGGCCACGGACAAAGTGGAGGGAGGCTCAA
GACTTGGAGAGACTCAGGAATCAACATTCTTCATCACCATTAAGGTGCTGCTGAAAATGAG
AAGCAGTATGTGGTGGAGCCAACTTGTGAAACAAGGCCGTGTCAGAACAAACGGCTTCTACTC
GCTCCACGTGCAAGACTGGTTGGCCTCCACAAGACCTGCAAGCCTCTGGTGAAGCAGGGCTGCG
ACACTGACCGCCTGCCCTGCAGCAAGACCTGCTGAACCTGGCTGACATTGGCTCGTCATCGAC
GGCTCCAGTGTGGGACGGCAACTCCGACCGCTCTCCAGTTGTGACCAACCTCACCAA
AGAGTTGAGATTCGACACGGACACGCGATCGGGCCGTGCAAGTACACCTACGAACAGCGGC
TGGAGTTGGGTCGACAAGTACAGCAGCAAGCCTGACATCCTCAACGCCATCAAGAGGGTGGC
TACTGGAGTGGTGGACCAAGCACGGGGCTGCCATCAACTCGCCCTGGAGCAGCTTCAAGAA
GTCCAAGCCAAACAAGAGGAAGTTAATGATCCTCATCACCGACGGAGGTCTACGACGACGCC
GGATCCAGCCATGGCTGCCATCTGAAGGGAGTGTACACCTATCGCATAGGCCTGGCTGGCT
GCCCAAGAGGAGCTAGAAGTCATTGCCACTACCCGCCAGAGACCCTACGAAACAGCGGC
GTTTGACAACCTCATCAGTATGTCCTCAGGATCATCCAGAACATTGTACAGAGTTCAACTCAC
AGCCTCGGAACTTGAATTCAAGAGCAGGCAGAGCACCAGCAAGTGTGCTTACTAATGACGTGTT
GGACCAACCCACCGCTTAATGGGCACGCCAGGTGCATCAAGTCTGGCAGGGCATGGAGAAC
AAATGTCCTGTTATTATTCTTGCCATCATGCTTTTCAATTCCAAAAGTGGAGTTACAAGA
TGATCACAAACGTATAGAATGAGCCAAAAGGCTACATCATGTTGAGGGTGTGGAGATTACAT
TTTGACAATTGTTCAAAATAATGTCGGAATACAGTGCAGGCCCTAACGACAGGCTTACGTAG
AGCTTTGTGAGATTGTTAAGTGTATTCTGATTGAACTCTGTAACCTCAGCAAGTTCA
TTTGTCATGACAATGTAGGAATTGCTGAAATTGTTAGAAGGATGAAAAAATAAAAAAA
AAAG

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRKPFTVPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVYGTDVASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR
ESFIVLESKPKKGTVTYPALTYSSSKPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA
TPPTLPRPSAASTTSIPRQSVGHRSQEMDLWSTATYTSSQNRPADPGIQRQDPSGAAFKQP
VGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATHFNLKHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTLQPLVKRVCDTRLACSKTCLNSADIGFVIDGSSVGTGNFRTVLQFVTN
LTKEFEISDTDTRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKRVGWWSGGTSTGAAINFALSQL
FKKS KPNKRKLMLITDGRSYDDVRI PAMA AHLKGVITYAIGVAWA AQEELEVIATHPARDHSFF
VDEFDNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCGTTAGGAGGTGGCTGCGTTGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGTTTCTGTTTCAGAGTAGTTACAACACAGATCTGAGTGTAAATTAAAGCATGGAAT
ACAGAAAACAACAAAAACTTAAGCTTAATTCTATCTGGAACTCCACAGTTCTAGCTCCCTGGACCC
GGTTGACCTGTTGGCTTCCCGCTGGCTCTATCACGTGGTGCCTCCGACTACTCACCCGAGTGTA
AAGAACCTTCGGCTCGCGTCTGAGCTGCTGTGGATGGCCTCGGCTCTGGACTGTCCTCCGAGTA
GGATGTCAGTGGAGATCCCTCAAATGGAGCCTCCTGCTGTCAGTCTGAGTTCTTGTGATGTGGTAC
CTCAGCCTCCCCACTACAATGTGATAGAACCGGTGAACTGGATGTAATTCTATGAGTATGCCGATTAA
CAGACAAGACTTCACCTCACACTCGAGAGCATTCAAACACTGCTCTCATCAAATCCATTCTGGTCATT
TGGTGACCTCCCACCCCTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTGGGTGAAAAAGTCT
TGGTGGGATATGAGGTTCTACATTTCATTAGGCCAAGAGGCTGAAAAGGAAGACAAAATGTTGC
ATTGTCCTTAGAGGATGAACACCTCTTATGGTGACATAATCGACAAGATTAGACACATATAATA
ACCTGACCTGAAAACCATTATGGCATTCAAGGTGGTAACTGAGTTGCCCCAATGCCAAGTACGTAAATG
AAGACAGACACTGATGTTTCACTCAACTGGCAATTAGTGAAGTATTTAAACCTAAACACTCAGA
GAAGTTTTCACAGGTATCCTCTAATTGATAATTCTCTATAGAGGATTACCAAAACCCATATTT
CTTACCAAGGAGTATCCTCAAGGTGTTCCCTCCACTGCAGTGGTTGGTTATATAATGCCAGAGAT
TTGGTGCCAAGGATCTATGAAATGATGGTCACGTAACCTAAACCATCAAGTTGAAGATGTTATGTCGGGAT
CTGTTGAATTATTAAAAGTGAACATTCACTCCAGAACAGACAAAATCTTCTTCTATAGAACATCC
ATTGGATGTCGTCAACTGAGACGGTGTGATTGCAGCCATGGCTTTCTCAAGGAGATCATCACTTT
TGGCAGGTCTGTAAGGAACACCATGCCATTAAACTTCACATTCTACAAAAGCCTAGAAGGACAG
GATACCTGTGAAAGTGTAAATAAAAGTAGGTACTGTGAAAATTCTATGGGAGGTCAAGTGTGCTGGCT
ACACTGAACACTCATGAAAACCCAGACTGGAGACTGGAGGGTTACACTGTGATTATTAGTCAGG
CCCTTCAAAGATGATGTGGAGGAATTAAATATAAGGAATTGGAGGTTTGCTAAAGAAATTAAATAGG
ACCAAACAATTGGACATGTCATTCTGTAGACTAGAATTCTAAAGGGTGTACTGAGTTATAAGCTCA
CTAGGCTGAAAAACAAATGTAGAGTTATTGAAACATGTAGTCATTGAAAGGTTGTGTA
TATCTTATGTGGATTACCAATTAAATATGTAGTGTCTGTGCAAAACTCTTCACTGAAGTTATA
CTGAACAAAATTTCACCTGTTGGTCATTATAAGTACTTCAGATGTCAGTGTGCAAGTGTGCTGGCT
ATTATTAAAATTACTCACTTGTGTTAAATGTTGACGATTCAACAGATAAAAGGATAG
TGAATCATTCTTACATGCAAACATTCCAGTTACTTAACAGTGTGCAAGTTATTGATACATCACTCCA
TTAATGTAAGTCATAGGTCAATTGATATCAGTAATCTCTGGACTTGTAAATATTACTGTGGT
AATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPIYRQDFHF
TLREHSNC SHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFFLLGQEAEKEDKMLA
LSLEDEHLLYGDIIIRQDFLDTYNNLTLKTI MAFRWVTEFCPNAKYVMKTDTDVFINTGNLVKYLL
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRDLVPRIYEMMGHV
KPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

FIGURE 38

MELGCWTQLGLTFLOQLLISSLPREYTGINEACPGAEWNIMCRECCYDQIECVCPGKREVVGYT
IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGTLDDFYVKGFYCAECRAGWYGGDCMRCGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGNDRGQII
KRCVGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSGNE
KRTCQQNGEWSGKQPIKACREP KISDLVRRRVLPMQVSRETPLHQLYSAAFSKQKLQSAPTK
KPALPFGDLPMGYQHLHTQLQYECISPFYRLGSSRTCLRTGKWSGRAPSCIPICGKIENITAP
KTQGLRPWPQAAIYRRRTSGVHDGSLHKGAFLVCGALVNERTVVVAHCVTDLGKVTMIKTADL
KVVLGKFYRDDDREKTIQSLQISAIILHPNYDPILL DADIAILKLLDKARI STRVQPICLAASR
DLSTS FQESHITVAGWNVLADVRSPGFKN DTLRSGVVSVVDSLLCEEQHEDHGIPVS VTDNMFCA
SWEPTAPS DICTAETGGIAAVSFPGRA SPEPRWHLMGLVWSYDKTCSHRLSTAF TKVLPFKDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTTCCATACCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCGTGATTATTAACGTGGCTTAATC
TGAAGGTTCTCAGTCAAATTCTTGTGATCTACTGATTGTGGGGCATGGCAAGGTTGCTTAAAGGAGCTTGGCTGG
TTTGGGCCCTGTAGCTGACAGAACGGTGGCCAGGGAGAATGCAGCACACTGCTCGAGAATGAAGGCCTCTGTG
TGGCTTGCCCTGGCTCAGCCTGCTAACATGACAATGTGGCAACCTGCACCTCCGTATTAGAACTCTGTA
AAGGTGCCCTCCACTACGGCCTGACCAAAGATAGGAAGAGGGCCTCACAGAGATGGCTGTCCAGACGCCGTGCGAGCC
TCACAGCCACGGCTCCCTCCCGAGGGTTCTGCAGCTGCCACCCTCCTTAATGACAGACGAGCCTGGCCTAGACA
ACCCCTGCCAACGTGTCTCGGCAGAGGACGGGCAGCCAGCAATCAGCCCAGTGGACTCTGGCCGGAGAACCGAACTA
GGGCACGGCCCTTGAGAGATCCACTATTAGAACGAGATCATTTAAAAAAATCGAGCTTGAGTGTCTCGAA
GGACAAAGAGCGGGAGTGCAGTGCACCATGCCGACCAGGGCAGGGAAAATTCTGAAAACACCACTGCCCTGAAG
TCTTTCAAGGGTGTACCACTGATTCCAGATGGTGAATTACAGCATCAAGATCAATCGAGTAGATCCCAGTGAAA
GCCCTCTATTAGGCTGGTGGAGGTAGCAGAACCCCCACTGGCCATATCATTATCCAACACATTATCGTGATGGGG
TGATGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTC
ACAACACTACGCTGTGCGTCTCTGCCAGCCCTGCCAGGTGCTGCTGACTGTGATGCGTAACAGAAAGTTCCGCA
GCAGGAACAATGGACAGGCCCGGATGCCAACAGACCCCGAGATGACAGCTTCATGTGATTCTCAACAAAAGTAGCC
CCGAGGAGCAGCTGGAATAAAACTGGTGCAGCAAGGTGGATGAGCCTGGGTTTCATCTTCAATGTGCTGGATGGCG
GTGTGGCATATCGACATGGTCACTGATTCAAGGCCAGTGAAGAACAGTGTTCACCTCGCTGTCCGCCAGGGTCGGCAGC
GCCCAAGGGCTGACATCTTCAGGAAGCCGCTGAAACAGCAATGGCAGCTGGTCCCCAGGGAGAGGAGCAACA
GGAGCCCTGACATCTTCAGGAAGCCGCTGAAACAGCAATGGCAGCTGGTCCCCAGGGAGAGGAGCAACA
CTCCCAAGCCCTCCATCTTACAATTACTGTGATGAGAAGGTGGTAAATATCCAAAAGACCCGGTAATCTCTCG
GCATGACCGTGCAGGGGGAGCATCACATAGAGAATGGGATTTGCCATCTATGTGATCATCAGTGTGAGGCCGGAGGAG
TCATAAGCAGAGATGGAAGAATAAAACAGGTGACATTGGTGAATGTGGATGGGTCGAACATGACAGAGGTCAAGC
GGAGTGAGGAGCTGGCATTATTGAAAAAGAACATCATCCTCGATGACTCAAGCTTGGAAAGTCAAAGAGTATGAGC
CCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACACATGGCCCCACCCAGTGAACGGTCCCCATCCT
GGGTGATGTGGCTGGAATTACCAACGGTGCCTGTATAACTGTAAAGATATTGTATTACGAAGAACACAGCTGGAAAGTC
TGGGCTTCTGCATTGTAGGGTCAAGAGAAAACAGAAAAATCACAATAGGCTAAGAAGTGAACATCAGGAATGA
CACAGCATACAATGATGGAAGAATTAGATGTGGTGAATTCTTCTGCTGTCATGGTAGAAGTACATCAGGAATGA
TACATGCTGCTGGCAAGACTGCTGAAAGAACCTAAAGGAAGAATTACTCTAACTATTGTTCTTGGCTGGCACTT
TTTATAGAATCAATGATGGTCAGAGAAAACAGAAAAATCACAATAGGCTAAGAAGTGAACACTATATTATC
TTGTGATTTTATATTAAAGAACATCATTGAAAAATGTCAGGAAAAGTATGATCATCTAATGAAAGCCAGTT
ACACCTCAGAAAATATGATTCAAAAAAAATTAAGAACACTAGTTTTTCTGAGTGTGGAGGATTCATTACTCTAC
AACATTGTTATATTCTATTCAATAAAAGCCCTAAAACAACAAATGATTGATTTGTATAACCCACTGAATT
CAAGCTGATTAAATTAAAATTGGTATATGCTGAAGTCTGCCAAGGGTACATTATGCCATTAAATTACAGCT
AAAATTTTAAATGCAATTGCTGAGAACGTTGCTTCAAAACAAGAATAATTTTCAAGAAGTTAA

FIGURE 40

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDCASLTATAPS
PEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDGRSNRTRARPFERSTIRSRSFKKINR
ALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDPSESLSIRLV
GGSETPLVHIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNGQAPDAYRPRDDS FHVILNKSSPEEQLGIKLVRKVDEPGVFI FNVLDDGVAYRHG
QLEENDRVLAILINGHDRLRYGSPEAAHLIQASERRVHLVVSRQRSPDIFQEAGWNSNGSWSPG
PGERSNTPKPLHPTITCHEKVNNI QKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISR DGR
IKTGDILLNVDGVELTEVRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP
SDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGT PAYNDG
RIRCGDILLAVNGRSTSGMIHA CLARLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAGGCATTGTATCTCAGTTGTCATCAAGTCGCAATCAGATTGGAAAAGCTCAACTGAAGCTT
CTTGCCTGCAGTGAAGCAGAGAGATAGATATTACCGTAATTCACGTAATAAAAACATGGGCTCACCTGACT
TTCCACCTTCCTACAAATTCCGATTACTGTTGCTGTTGACTTGCCTGACAGTGGTTGGTGGC
CACCACTGAGGAAAGGGAAAAACTGACTAATGAAGCATCCACGAAGAAGGTAGAACCT
ATAAGACCCCTCATTTGGGAAAGGGAAAAACTGACTAATGAAGCATCCACGAAGAAGGTAGAACCT
GACAACGTGCTCTGTCCTACCTCAGAGGCCAGAGCAAGCTCATTTCAAACCAGATCTCAC
TTTGGAAAGAGGTACAGGCAGAAAATCCAAAGTGTCCAGAGGCCGGTATGCCCTCAGGAATGAAAG
CTTACAGAGGTGCCATCCTCGTCCCCACCGAACAGAGAGAACACCTGATGTACCTGCTGGAA
CATCTGCATCCCTCTGCAAGGCCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGG
TAAAAGTTAATCGAGCAAACCTCTGAATGTGGCTATCTAGAACGCTCAAGGAAGAAAATTGGG
ACTGCTTATATTCCACGATGTGGACCTGGTACCCGAGAACAGTACTTTAACCTTACAAGTGTGAGGAG
CATCCCAAGCATCTGGTGGTGGCAGAACAGCACTGGTACAGGTTACGTTACAGTGGATATTGG
GGGTGTTACTGCCCTAACGAGAGCAGTTTCAAGGTGAATGGATTCTCTAACAACTACTGGGAT
GGGGAGGCGAACAGCATGACCTCAGACTCAGGGTGAGCTCCAAAGAACATGAAAATTCCGGCCCTG
CCTGAAGTGGTAAATATAACATGGCTTCCACACTAGAGAACAGCAATGAGGTGAACGCAGAACG
GATGAAGCTTACACCAAGTGTACAGCTGGAGAACAGATGGGTGAGTAGTTGTTCTATAAAAT
TAGTATCTGTGGAACACAATCCTTATATACACATCACAGTGGATTCTGGTTGGTGCATGACCC
TGGATCTTGGTGTGTTGGAAGAACACTGATCTTGTGCAATAATTGGCTAGAGACTTCAA
ATAGTAGCACACATTAAGAACCTTACAGCTCATGTTGAGCTGAATTTCCTTTGTATTTCT
TAGCAGAGCTCCTGGTGTAGAGTATAAACAGTTGTAACAAGACAGCTTCTAGTCATTTGAT
CATGAGGGTTAAATATTGTAATATGGAACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAA
AATGAACGCTATTGAGGACTCTGGTGAAGGAGATTATTTAAATTGAGTAATATATTATGGGAT
AAAAGGCCACAGGAAATAAGACTGCTGAATGTCAGAGAACAGAGTTGTTCTCGTCCAAGGTAGAA
AGGTACGAAGATAACAATACTGTTATTCACTTACGTCAGGAGAACAGAGTTGTTCTCGTCCAAGGTAGAA
GAGAAGGGTCCACAAAGAGGGAGAAAAGGCAGCAATCAGGACACAGTGAACCTGGGAAATGAAGA
GGTAGCAGGAGGGTGGAGTGTGCGCTGCAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGC
CTTCAGGGGAGGACCTGCCAGGTATGCCCTCAGTGTGAGCCCACCAAGAGAACATCATTCTCTATTAGT
TTTAAAGAGTTTGTAAAATGATTTGTACAAGTAGGATATGAATTAGCAGTTACAAGTACAT
ATTAACATAATAAAATATGTCTACAAATACCTGAGTAAAGTGTGAAAAAGCAAA

FIGURE 42

MGFNLTFHLSYKFRLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLLGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQOLDYGIYVIHQAEGKKFNRAKLLNVGYLEALKEENWDCFIFHDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRSGYFGGTALSREQFFKVNGFSNNYWGWGGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRDGSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGATGGACTGAGCTCCCAGATCTGGG
CCGCTTGCCCTCGCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTCCCACAACAG
ACGGGACAACTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCAT
GTTCCAGAGGCAGAAGGAGGCAGACACCCACTTCCCCATCTGCATTTCTGCTGCGGCTGTC
ATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCGTCCCTCCC
TTCCTTATTATTCCCTGCTGCCAGAACATAGGTCTTGGAAATAAAATGGCTGGTCTTTGTTT
TCCAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 44

MALSSQIWAACLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRDTHFPI
CIFCCGCCHRSKCGMCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTCCCCAACATGCCTCACCC
TCATCTATATCCTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT
TCCGTTGGTGGGCCGTGACTTCCCCCTGAAGTCAAAGTAAGCAAGTTGACTCTATTGTCTG
GACCTCAACACAACCCCTTGTCAACCATAAGCCAGAAGGGGGCACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACGTGAAG
AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCTAAAGTCACCATGGGCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGATGGAACATGGGAAGAGGATGTGATT
TATACTGGAAGGCCCTGGGCAAGCAGCCAATGAGTCCATAATGGTCCATCCTCCCCATCTC
CTGGAGATGGGAGAAAGTGTATGACCTCATCTGCCTGCCAGGAACCCGTCAAGCAGAAACT
TCTCAAGCCCCATCCTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATG
GTCCTCCTGTGTCTCTGTTGGTCCCCCTCTGCTCAGTCTCTTGACTGGGCTATTCTTG
GTTTCTGAAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTGGGG
AAACTCTAACATATGCCCTTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAACTCTAACAGGAAAGATCCAGCAAATCGGTTACTCCACTGTGGAAATACCGAAAAAGAT
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTGCCTATGAGAATGTTA
TCTAGACAGCAGTGCACTCCCCCTAAGTCTCTGCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAAAGPVKELVGSGGAVTFPLSKVKQVDSIVWTFTNTPLVTIQP
EGGTIIVTQNRNRERVDGPDDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLCCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLCLLLVPLLLSSLFVLGLFLWFLKRERQEEYIE
EKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCCTTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTCTAAAACCCATCTCTGCTTGAGTGGTGGTCCCAGGAATT
ATAGGAGCAGGTCTGATGCCATTCCAGCAACAAAATGTCCTGACAGCAAGAAAAAGAGCGTG
CTGCAACAACAGAACAGAACTGGAATGTTCTTCATCATTTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTTAAAGGTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGAAATTTCATTGAAAAACATCAGTGACATTCACCAGAATCCTCAA
CTTGCAGTGGTTTCCAATGACTCTGTGCACCCTACTGGTTCAATAAAACCCACCACTGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTCCACTTCGATTCTGAAAGAAAACAAACATAGG
CTTATCCACTTCCACTATTAGGTCTATTGCTTGGAAATTCTGGAGGTCTGTTGGCT
CAGTCAGATAGTCATCGGTTCCCTGGCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG
TGTATTAATGGGAATAAAATGTAAGTATCAGTAGTTGAAAAAA

FIGURE 48

MTCCEGWTSCNGFSLVLLLGVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNANCEFSL
KNISDIHPESFNQLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHFSVFLGL
LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGGCTGCAGACCCCCCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCTCGAACTGTGA
CATGGAGAGAGTGACCCCTGGCCCTCTCCTACTGGCAGCCTGACTGCCTTGGAAAGCCAATGACC
CATTTGCCAATAAGACGATCCCTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGC GGAGGGCTCCTGGCATTGCTGGATCGCGGAGTTCTGAGTGCAAATGCAAATACAAGAG
CAGCCAGAACGACAGCACAGTCCTGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTCTGCCA
CTACTTGCTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACC
TCCTCCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCC
TTCTGATCAGGAGGCTTCTTATGAATTAAACTCGCCCCACCACCCCTCA

FIGURE 50

MERVTLALLLLAGLTALLEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCKYKS
SQKQHSPVPEKAIPILTPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCAGGCAGTTGAGGAAGGACAGGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGCAAGG
AGGAGACCCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGCTGGCAGAGAATGAAGTTCCAG
GGGCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGGAGTGGGAGGCTGGCCCTGCAGAG
CGGAGAGGAAAGCAGTGGACAAATATTGGGGAGGCCCTGGACATGGCCTGGGAGACGCCCTGA
GCGAAGGGTGGAAAGGCATTGGCAAAGAGGCCGGAGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTGCCAAGGGACCAGAGAACAGTGGCACTGGAGTCAGGCAGGTTCCAGGCTTGGC
AGCAGATGCTTGGCAACAGGGTCGGGAAAGCAGCCCATGCTCTGGGAAACACTGGGACAGA
TTGGCAGACAGGCAGAAGATGTCATTGACACGGAGCAGATGCTGTCCGGCTCTGGCAGGG
GTGCGCTGCCACAGTGGTGCCTGGGAAACTTCTGGAGGCATGGCATCTTGGCTCTCAAGGTGG
CCTTGGAGGCCAGGGCAAGGGCAATCCTGGAGGTCTGGGACTCCGTGGTCCACGGATAACCCG
GAAACTCAGCAGGCAGCTTGGATGAATCCTCAGGGAGCTCCCTGGGTCAGGAGGCAATGGA
GGGCCACCAAACCTTGGACCAACACTCAGGGAGCTGTGGCCAGCCTGGCTATGGTCAGTGAG
AGCCAGCAACCAGAAATGAAGGGTCACGAATCCCCCACCACCTGGCTCAGGTGGAGGCTCCAGCA
ACTCTGGGGAGGCAGCGGCTCACAGTCGGCAGCAGTGGCAGCAATGGTACAACAAC
AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGTGGCAGCAG
TGGCGGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGAGTGA
CCTCCTGGGATCCAGCACCGGCTCCTCCGGCAACCACGGTGGGAGCGGGAGGAAATGGGATTCA
CATAAACCCGGGTGTGAAAAGCCAGGGAAATGAAGCCCAGGGAAATCTGGGATTCA
CTTCAGAGGACAGGGAGTTCCAGCAACATGAGGGAAATAAGCAAAGAGGGCAATGCCCTTG
GAGGCTCTGGAGACAATTATCGGGGCAAGGGTCAGCTGGGAGCAGTGGGAGGTGACGCTGTT
GGTGGAGTCATACTGTGAECTGAGACGTCCTGGATGTTAACTTGAACACTTCTGGAA
GAATTTAAATCCAAGCTGGTTCATCAACTGGGATGCCATAAACAGGACCAGAGAAGCTCTC
GCATCCCGTGACCTCCAGACAAGGAGCCACAGATTGGATGGGAGGCCACACTCCCTCTAA
AACACCACCCCTCTCATCACTAATCTAGCCCTGCCCTGAAATAACCTTAGCTGCCCAACAAA
AA
AAAAAAAAAAAAAAAAAAAAA

FIGURE 52

MKFQGPLACLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGAAGSKVEALGQGTREAVGTGVQRQVPGFGAADALGNRVEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGVPGHSGAWETSGGHGIFGSQGGLGGQQGNPGGLGTPWVHGYPGNSAGSGMNPQGAPWGQGGNGGPPNFGTNTQGAVAQPGYGVSRASNQNEGCTNPPSGGGSSNSGGSGSQGSSGSGSNGDNNNGSSSGSSGSSGSSGSSGGSSGGSSGNSGGSGDGSSESSWGSSTGSSSGNHGGGGNGHPGCEKPGNEARGSGESGIQGFRGQGVSSNMREISKEGNRLGGSGDNYRGQGSSWGSGGDAVGGVNTVSETSPGMFNFDTFWKNFKSKLGFINWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80, 90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161, 159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224, 236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252, 253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285, 283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301, 298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329, 325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389, 387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGGGACAAGCTGCTCCGACAGAAGGGATGTCGCTGCTGAGCCTGCCCTGG
CTGGGCCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGCTGGTTGTGGCTCCTGGCT
ACTCGCCCGCATCCTGGCTGGACCTATGCCCTCTATAACAAACTGCCGCCGCTCCAGTGTTC
CACAGCCCCAAAACGGAACTGGTTTGGGTCACCTGGGCTGATCACTCCTACAGAGGAGGGC
TTGAAGGACTCGACCCAGATGTCGGCACCTATTCCCAGGGCTTACGGTATGGCTGGGCTCCAT
CATCCCCCTTCATCGTTTATGCCACCTGACACCACCGGCTATCACCAATGCCCTAGCTGCCA
TTGCACCCAAGGATAATCTCTTCATCAGGTTCTGAAGCCCTGGCTGGAGAAGGGACTGCTG
AGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGACGCCGCCCTCCATTCAACATCCT
GAAGTCCTATATAACGATCTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGG
CCTCAGAGGGCAGCAGTCGCTGGACATGTTGAGCACATCAGCCTCATGACCTGGACAGTCTA
CAGAAATGCATCTCAGCTTGACAGCCATTGTCAGGAGAGGCCAGTGAATATATTGCCACCAT
CTTGGAGGCTCAGTGCCCTGTAGAGAAAAGAACGCCAGCATATCCTCCAGCACATGGACTTCTGT
ATTACCTCTCCATGACGGCGCGCTTCCACAGGGCTGCCCTGGTGCATGACTTCACAGAC
GCTGTCATCCGGGAGCGCGCTCGCACCCCTCCCCACTCAGGGTATTGATGATTTTCAAAGACAA
AGCCAAGTCCAAGACTTGGATTTCATTGATGTGCTTGCTGAGCAAGGGATGAAGATGGGAAGG
CATTGTCAGATGAGGATAAAGAGCAGAGGCTGACACCTCATGTTGGAGGCCATGACACCACG
GCCAGTGGCCTCTCCTGGGTCTGTACAACCTTGCAGGGCACCAGAAATACCAGGAGCGCTGCCG
ACAGGAGGTGCAAGAGCTCTGAAGGACCCGATCTAAAGAGATTGAATGGGACGACCTGGCCC
AGCTGCCCTCCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTCATC
TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCGAGTCATCCCCAAGGCATTACCTG
CCTCATGATATTATAGGGTCCATCACAACCCACTGTGTTGGCCGATCCTGAGGTCTACGACC
CCTTCCGCTTGACCCAGAGAACAGCAAGGGAGGTACCTCTGGCTTTATTCTTCTCCGCA
GGGCCCAGGAACTGCATGGGAGGCAGGCGTTGCCATGGGGAGATGAAAGTGGTCTGGCGTTGAT
GCTGCTGCACTTCCGGTCTGCCAGACCAACTGAGCCCCGCAAGGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTGGCTGCCAGGCTGAATGTAGGCTTGCAGACTTCTGAC
CCATCCACCTGTTTTGCAAGATTGTCATGAAATAACGGTGCTGTCAAA

FIGURE 54

MSLLSLPWILGLRPVAMSPWLLLLLVGSWLLARILAWTYAFYNNCRLQCFQPQPKRNWFWGHLG
LITPTEEGLKDSTQMSATYSQGFTVWLGPPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
WLGEGLLSGGDKWSRHRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASSEGSSRLDMFEHI
SMLTLDLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHGRRFHAC
RLVHDFTDAVIRERRTLPTQGIDDDFKDKAKSKTLDFIDVLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLARHPEYQERCRCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPVEVDPFRDPENSKGSP
LAFIPFSAGPRNCIGQAFAEMKVVLALMILLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTGA
GCCTACTCGTTGATTGCAACTATCATGGTGTGTTGCACTTACCCGTGTTCTGCCT
TTGGTGGATAACAAGGGACTGCACTTATCTTCTGCATTGGCAGTCTTGGCATTGACGTGG
TACAGCCTTCCTTCATACCAATTGCAAGGGATGCTGTGAAGAAGTGTGCGGTGTCTGC
ATAATTCATGCCAGTTATGAAGCTTGAAAGGCACATGGACAGAAGCTGGTGGACAGTTT
GTAACATCTCGAACACCTCTGTCTTACAGACATGTGCCTTATCTTGCAGCAATGTGTTGCTT
GTGATTGAAACATTGAGGGTTACTTTGAAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
GCACAGGATGAGAAGTGGGTTCTGTATCTGTGGAGTGGAAATCTCCTCATGTACCTGTTCTC
TCTGGATGTTGTCCCACTGAATTCCCATGAATACAAACCTATTCAACAGCAACAGCAAAAAAAA
AAA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCASFWWHNKGALIFCILQSLALTWYSLSFIPFAR
DAVKKCFAVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCAGTGGAGCACCCAGCAGGCCAACATGCTCTGTCTGTGCCTG
TACGTGCCGTATCGGGGAAGCCCAGACCGAGTTCCAGTACTTGTAGTCGAAGGGCTCCCTGCCAGCTGAAGTCC
ATTTCAAGCTCACTGTTCATCCCCCTCCAGGAATCTCCACCTACCGCCAGTGGAGCAGAAAATTGTACAAGCT
GGAGATAAGGACCTTGATGGCAGCTAGACTTGAAGAATTGTCATTATCTCAAGATCATGAGAAGAAAGCTGAGG
CTGGTGTAAAGATTGGACAAAAGAATGATGGACGATTGACGCCAGGAGATCATGCAGTCCCTGCCGGACTTG
GGAGTCAAGATACTGAACAGCAGCAGAAAAATTCTCAAGAGCATGGATAAAACGCCACGATGACCATCGACTGG
AACGAGTGGAGAGACTACCACCTCCACCCGTGGAAAACATCCCGAGATCATCCTACTGGAAGGATTCCACG
ATCTTGTGTGGTGTGAGAATCTAACGGTCCCGATGAGTTACAGTGGAGGAGGGAGACGGGATGTGGTGGAGA
CACCTGGTGCAGGAGGGTGGGGCAGGGCCATTCACGAACCTGCACGCCCTGGACAGGCTCAAGGTGCTCATG
CAGGCTCATGCCCTCCGCAGCAACACATGGGCATCGTGGCTTCACTCAGATGATTGAGAAGGGAGGGCCAGG
TCACTCTGGCGGGCCAATGGCATCACGCTCTCAAAATTGCCCTGAGATGCCATCAATTATGCCATGAGCAG
ATCAAGGCCCTTGTGGTAGTGACCAGGAGACTCTGAGGATTACAGAGGGCTTGGCAGGGCTTGGCAGGGCC
ATGCCCGAGCACCATCACCAATGGAGGCTCTGAGACCCGGATGGCGTGGGAAGACAGGCCAGTACTCAGGA
ATGCTGGACTGCCAGGAGGATCTGGCCAGAGGGGGTGGCCCTTACAAAGGCTATGTCCTTACATGCTG
GGCATCATCCCTATGCCATCGACCTGAGCTCACAGCCTGGCTTGAGCAGTGGCAGACTATGCCAGT
AACAGCGGGACCCGGCTTGTGTGCTCTGGCTTGACGGCAGCTTCAACTGTGGCAGCTGGCAGCTAC
CCCCCTGGCCCTAGTCAGGACCCGGATGCAGGCCAAGCCTTATTGAGGGCCTCCGGAGGTGACCATGAGCAGCCTC
TTCAAACATATCTCGCGACCGAGGGGCCCTCGGGCTGTACAGGGGCTGGCCCCCAACTTATGAGGCTCATCCA
GCTGTGAGCCTACGCTACGTGAGAACCTGAAGATCACCCCTGGCGTGCAGTGGGGTGTGACGGCCAGGGC
CGCCGGCAGTGGACTCGCTGATCTGGCGCAGCCTGGGGTGTGACGGCCATCTCATCTGTGAATGTGCCAACACT
AAGCTGTCTCGAGCCAAGCTGTGAAACCCCTAGACGCCACCCGAGGGAGGGTGGGGAGGCTGGCAGGCCAGGGT
GTCTGCTGACCCAGCAGACCCCTCTGGCTTCCAGCAGAACACAGGCACTTCTAGGGCTCAGGGTCAAGCAG
CTCCGGGCTCACATGTGTAAAGGACAGGAACTTCTGCTGAGCTGGCTGCCAATGAGCTGGAGGCTGGAGGCC
TAGTTCTTCCATTCTCACCTTGAGCAGCTGTTGGCACGGGGCTCTGCTGCGTGCATCTCCCTGTGC
CCTCTGCTGCTGCTGCTGCTGAGGTAAGGTGGAGGGAGGCTACAGCCCCACATCCCCACCCCTGTCCAATCCC
ATAATCCATGATGAAAGGTGAGGTACGTGGCTCCAGGCCCTGACTTCCCACCTACAGCATTGACGCCAACITGG
TGTGAAGGAAGGAAAGGATCTGGCTTGTGGTACTGGCATCTGAGCCCTGCTGATGGCTGGGCTCTCGGGCATG
CTGGGACTGCAAGGGGCTGGGCTGGCTGGCTGACAGAACGCAAGTGTGGCTCATGGTGTCTGAGCT
GGCCTGGACCTGCTGAGATGGGCCCCACCTCAGAACCAAACACTGTGCCCCACTGTGGCATGAGGGCAGTGGAGCA
CCATGTTGAGGGCGAAGGGCAGACGCTTGTGTCTGGGGAGGAAGGAAAGGTGGAGGCCATTATAGG
ACTGTTGGGAAAGGGTTTGTCCAGAAGGACAAGCCGACAATGAGCAGCTTGTGCTTCCAGAGGAAGACGAGG
GAGCAGGAGCTGGCTGACTGCTCAGAGTCTGTTGACGCCCTGGGGTCTGTCACCCAGCAGGGCGCAG
GGGACCCACATTCCACTTGTGACTGCTTGGGACCTTATTATTTGATTATTTGAAACAGGTTATGTCT
AACTTTTATGATTGTTAATTATAGCTGTCTTCAAGTCTTATTTTATTATGTTCATGTTGTT
GATTGACCTCCAAAGGCCCTGGGATGGGAGGGAGGAGAAGGGGGCTTGGGCCCTGCACTCACACT
GTCCAGAGAAATTCTTGGACTGGAGGAGCAGAAAGCAGGCCAGAAGGAGCAGGCCCTGGCTCTTCTTGGAG
GTGGGGAGGGCTGCCCTGGCTTAGGATTTCAGGGTTTACTGGGCTGGAGAGAGAGGGAGGAACCTCAAT
AACCTGAAAGGTGAAATCCAGTTTCTGCTGCGAGGGTTCTTATTTCACTTCTGAAATGTCAGGAG
TGAGGTGCTCTCACTGTGAAATTCTGTTGGGGGGCTGGAGGAGGGGTGGGCTCCGCTCCCTCCAGC
CTTCTGCTGCCCTTGCTAACATGCCGCCACTGCCACCTCACCTGGACTTCACTTCCATTCAACAGAATGACCTGA
TGAGGAAATCTCAATAGGATGCAAAGATCAATGCAAAATTGTTATATGAAACATATAACTGGAGTGTCAAAAAG
CAAATTAAAGAAAGATTGGACGTTAGAAGTGTCTTAAAGCAGCCTCTAATAAAGTGTGTTCAAGGCTGAAAAAA
AAA

FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFTYRQWKQKIVQAGDKLDG
QLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILKSMDKNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTEERQTGMWWRHLVAGGG
AGAVSRTCTAPLDLKVLMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIK
FMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIAQSSSIYPMEVLKTRMALRKTGQYSGMLDCARR
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNAWLQHYAVNSADPGVFVLLACGTMSSTC
GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCGAGATACGCTGGGAACCTTCCCCAGCCATGGC
TTCCCTGGGGCAGATCCTCTTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTGATTTAGGGAGACACTCCATCACAGTCACACTGTGCCTCAGCT
GGGAACATTGGGGAGGGATGGAATCCTGAGCTGCACTTTGAAACCTGACATCAAACTTCTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTTAGGCTTGGCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTCGGAGCAGGATGAAATGTTAGGCTGAGCAGTGTGATCAAGTGATAGTT
GGCAATGCCCTTTGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAAACTGGAGCCTTCAGCATGCCG
AAAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTGGGTGTGAGGCTCCCCGATGGTCCCC
CAGCCCACAGTGGCTGGCATCCAAAGTTGACCAGCAGGAGCCAACCTCTCGGAAGTCTCCAATAC
CAGCTTGAGCTGAACCTGAGAATGTGACCATGAAGGTTGTCTGTCTACAATGTTACGA
TCAACACACATACTCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGATATCAAAGTG
ACAGAATCGGAGATCAAAGGGAGTCACCTACAGCTGCTAAACTCAAAGGTTCTGTGT
CTCTTCTTCTTGCCATCAGCTGGCACTCTGCCTCTAGCCCTACCTGATGCTAAATAAT
GTGCCCTGGCCACAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTCAC
CACAGATATGACCTAGTTTATTTCTGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG
AGCAAACAAGAGCAAGAAAACAAAAAGCAAGCCAAAGCAGAAGGCTCCAATATGAACAAGATAAAT
CTATCTCAAAGACATATTAGAAGTTGGAAAATAATTGATGAACTAGACAAGTGTGTTAAGA
GTGATAAGTAAATGCACTGGAGACAAGTGCATCCACAGTCAGGGACCTCCCCCTGCCTGT
CACCTGGGAGTGAGAGGACAGGATAGTGCATGTTCTGTCTGAATTAGTTATATGTC
TGTAATGTTGCTCTGAGGAAGCCCTGGAAAGTCTATCCAAACATATCCACATCTTATATTCCAC
AAATTAAGCTGTAGTATGTAACCTAACGCGTCTAAATTGACTGCCACTTCGCAACTCAGGGCG
GCTGCATTTAGTAATGGGTCAAATGATTCACTTTATGATGCTTCAAAGGTGCCTTGGCTTC
CTTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTAGCATAAACAGAGCAGT
CGGGGACACCGATTTATAAAACTGAGCACCTCTTTAAACAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 6o

MASLGQILFWSIISIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKS
DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAGTYKC
YIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVSVLYNVTINNTYSCMIENDIAKATGDIKVTESEIKRRSHLQLLNSKASL
CVSSFFAISWALLPLSPYLMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTAGAATCACCATGGCCAGCTATCCTTACCGCAGGGCTGCCAGGAGCTGCAGGACAAG
CACCAAGGAGCCCCCTCGGGTAGCTACTACCCCTGGACCCCCAATAGTGGAGGGCAGTATGGTAGT
GGGCTACCCCTGGGGTTATGGGGTCTGCCCTGGAGGGCCTATGGACCACCAGCTGG
TGGAGGGCCCTATGGACACCCCAATCCTGGATGTTCCCTCTGGA~~ACTCCAGGAGGACCATA~~TG
GCGGTGCAGCTCCCCGGGGCCCTATGGTCAGCCACCTCCAAGTCTACGGTGCCAGCAGCCT
GGGCTTATGGACAGGGTGGGCCCTCCAA~~TGTGGATCCTGAGGCCTACTCCTGGTTCCAGTC~~
GGTGGACTCAGATCACAGTGGCTATATCTCATGAAGGAGCTAAAGCAGGCC~~TGGTCAACTGCA~~
ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTGACAAGACCAAGTCA
GGCCGCATCGATGTCTACGGCTTCAGCC~~GTGGAAATTCCAGCAGTGGAAGAACCTTT~~
CCAGCAGTATGACGGGACCGCTCGGCTCATTAGCTACACAGAGCTGCAGCAAGCTCTGCCC
AAATGGGCTACAACCTGAGCCCCAGTTCACCCAGCTTCTGGTCTCCGCTACTGCCACGCTCT
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA
GGCCTCCGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTCGAGGACTTCGTCA
CCATGACAGCTTCTCGGATGCTTGACCCAACCATCTG~~GGAGAGTGGAGTGACCCAGGGACCTT~~
TCCTGGCTTCTTAGAGTGAGAGAAGTATG~~GGACATCTTCTTTCC~~TGTCCCTCTAGAAGAAC
ATTCTCCCTGCTTGATGCAACACTGTTCAAAGAGGGTGGAGAGTCCTGCATCATGCCACCA
AATAGTGAGGACCGGGCTGAGGCCACACAGATAGGGCCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCTGTAATGG
AGTTAGTGTCCAGTCA~~GCTGAGCTCCACCC~~TGATGCCAGTGGTGAGTGTTCATCGGCTGTTACC
GTTAGTACCTGTGTTCCCTCACCAAGGCCATCCTGTCAAACGAGGCCATTTC~~TCTCCAAAGTGGAA~~
CTGACCAAGCATGAGAGAGATCTGCTATGGGACCAGTGGCTGGATTCTGCCACACCCATAAAT
CCTTGTGTGTTAACTCTAGCTGCC~~TGGGCTGGCC~~CTGCTCAGACAAATCTGCTCC~~TGGG~~CAT
CTTGCCAGGCTCTGCCCTGCAGCTGGACCCCTACTGCC~~TGCA~~GTCTGCTCGGCT
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTAATTG~~CATTTTTTC~~
ATTGGGCCAAAGTCCAGTGAAATTGTAAGCTCAATAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSTPGGPYGGAAPGGPYGQPPPSSYGAQQPGLYQGGAPPNDPEAYSWFQSVDSDH
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTGRIDVYGFSAWKFIQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCCTCGGCTCTGGCTGGTCTGTCTTCATC
TCCCAGGCCCTTTGCCCGAGCATGGTGTGGAGGAGAAAGTTCCAAAACCTCGGGACC
AACTTGCCCTAGCTCGGACAACCTTCTCCACTGGCCCTCTAACCTGAACATCCGCAGCCGC
TCTGGACCCTAGGTCTAATGACTGGCAAGGGTCCCTGAAGCTCAGCGTGCTCCATCAGATG
GCTTCCCACCTGCAGGAGGTCTGCAGTGCAGAGGTGGCCTCATCGTGGGCTGCCTGCCATG
GATTCCCTGGCCCCCTGAGGATCCTGGCAGATGATGGCTGCTGCCGTGAGGACCGCTGGGGA
AGCGCTGCCTGAAGAACTCTTACCTCTCAGTGCCTGCCCTCGCTCCGGCAGTGGCCCTT
TGCCTGGGAGTCTTCTCCCGATGCCACAGGCCCTCACCTGAGGCTTCACTCCTCCACCAGGAC
TCGGAGTCCAGACGACTGCCCGTTCTAATTCACTGGGAGCCGGGGAAAAATCCTTCCCAACG
CCCTCCCTGGCTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGTACCTGAATCCCAGTG
TGTCCCTGGGAGGTGGAGGCCCTGGACTGGTGGGAACGAGGCCATGCCACACCTGAGGGA
ATCTGGGGTATCAATAATCAACCCCCAGGTACAGCTGGGAAATATTAATCGGTATCCAGGAGG
CAGCTGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGAAATATTAATCGGTATCCAGGAG
GCAGCTGGGAAATATTCACTATACCCAGGTATCAATAACCCATTCCCTGGAGTTCTCCGC
CCTCCCTGGCTCTTCTTGGAACATCCCAGCTGGCTCCCTAACCTCCAAGCCCTAGGTTGCAGTG
GGGTAGAGCACGATAGAGGGAAACCCAACATTGGAGTTAGAGTCCTGCTCCGCCCTGCTG
TGTGGGCTCAATCCAGGCCCTGTTAACATGTTCCAGCACTATCCCCACTTTCACTGCCTCCCC
TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

FIGURE 64

MQGRVAGSCAPLGLLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQPAL
DPRSNDILARVPLKLSVPPSDGFPAGGSAVQRWPPSWGLPAMDSWPPEPDWPQMMAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRSNSLGAGGKILSQRP
PWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRPMPHPEGIWGINNQPPGTSWGNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWNIHLYPGINNPFPVGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCATCCCAGGAGCGCAGTGGCCACTATGGGTC
TGGGCTGCCCTTGTCCTCTTGACCCCTTGGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTGCAACTGAAGCTGAAGGAGTCTTCTGACAAATTCCCTATGAGTCCAGCTCCTGGAA
TTGCTTGAAAAGCTCTGCCTCCTCCATCTCCCTCAGGGACCAGCGTCACCCCTCACCATGC
AAGATCTCAACACCATGTTGTCTGCAACACATTGACAGCCATTGAAGCCTGTGTCCTTCTGGCCC
GGGCTTTGGGCCGGGGATGCAGGGAGGCAGGCCCGACCCGTCTTCAGCAGGCCACCC
CTGAGTGGCAATAAATAAATTGGTATGCTG

FIGURE 66

MGSGLPLVLLTLLGSSHGTGPGMTLQLKLKESFLTNSSYESSFLELLEKLCLLLHLPSGTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCGGC
CAGGTGCCCGTCGCAGGTGCCCTGGCCGGAGATGCGTAGGAGGGCGAGCGCGAGAACCCCC
TTCCTCGGCCTGCCAACCCGCCACCCAGCCCATGGCGAACCCGGCTGGGCTGCTTCTGGCG
CTGGGCCTGCCGTTCTGCTGGCCCGCTGGGGCGAGCCTGGGGCAAATACAGACCACTCTGC
AAATGAGAATAGCACTGTTGCCTCATCCACCAGCTCCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTCTCCCTCTGGCTGCCTTGCTCCTGGCTGTGGGCTG
GCACTGTTGGTGCAGGAAAGCTTCGGAGAACGCCAGACGGAGGGACCTACCGGCCAGTAGCGA
GGAGCAGTTCTCCCATGCAGCCGAGGCCGGCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCCATTAGGTCCCTCTGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGAAGAAGGTACTCAA
AGACTCTGCCCTGAGGTCAAGAGAGGATGGGCTATTCACTTTATATATTATAAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLALGLPFLARWGRAWGQIQTTSANENSTVLPSSSSDGTLRPEAITAIIVVS
LLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFShAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

FIGURE 70

MGLFRGFVFLLVLCLLHQSNSTFIKLNNNNGFEDIVIVIDPSVPEDEKIEQIEDMVTTASTYLFE
ATEKRFFFKNVSILIPEWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY
IHFTPDLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQFYRAKSKKIEATRCSAGISGRN
RVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDKVQTEKASIMFMQSIDSVVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPPVSLLKISQRIVCLVLDKGSMGGKDRLNR
MNQAAKHFLQLQTENGWSWGMVHFDFSTATIWNKLIQIKSSDERNTILAGLPTYPLGGTSICSGIK
YAFQVIGELHSQDGSEVLLTDGEDNTASSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG
SHFYVSDEAQNNGLIDAEGALTSGNTDLSQKSLQLESKGTLNSNAWMNDTVIIDSTVGKDTFFL
ITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSR
AANSSVPPITVNAKMNKDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA
GADSFKNDGVYSRYFTAYTENGRYSLKVRAGGANTARLKLRRPLNRAAYIPGWVVNGEIEANPP
RPEIDEDTQTTLEDFSRTASGGAFVVSQVPSLPLPDQYPPSQITDLDATVHEDKIIILTWTAPGDN
FDVGVKVQRYIIRISASILDLRDSFDDALQVNNTDLSPKEANSKESFAFKPENISEENATHIFIAI
KSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPTPTPTPTPDKSHNSGVNISTLVLSVIGSVVI
VNFILESTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGGAAACCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACTCCCCGGGCAGGGTGA
CAACAGGTGTCATCTTTTGATCTCGTGTGGCTGCCCTCTATTCAAGGAAAAGACGCAAGGTAACTTGA
GAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTAACCTTCCCTTCTGAAACCCCACTTGA
ACTCAACCAGCAAGCGGCCCTTCGGCTTAACCTTGTGGTGGAGGAGAGAACCTTGTGGGCTGCCCTCTTAGCA
GTGCTCAGAAGTGACTCCCTGAGGGTGACCAGAAGAAAGGAAGGTCCCCTCTGCTGTGGCTGCACATCAGGA
GGCTGTGATGGGAATGAAGGTGAAAATTGAGGATTTCACTTCAGTCATTGCTTCTGCCTGCAAGATCATCCTTAAA
AGTAGAGAAGCTGCTCTGTGGTGGTAACCTCAGGAGGAGAAGCTCGTCTAGAAGGAAATGGATGCAAGCACCTC
CGGGGCCCCAAAGCAGCTCTCTGGTGGCTAGCCCAGGGAAGCCCTGGGGGCCGGCTTGAGGGATGCC
ACCGGTTCTGGACGCACTGGCTGATTCTGAATGATGTGGTCTGGGGGGCTGCTGCTGGTGGATTCCCAGGGTGGT
GTTTGCTGGTCTCTCTGCTGTGCTATCTCTGTACATGTGGCTGCACCCAAAAGGTGACGAGGAGCAG
CTGGCACTGCCACGGGCAACAGCCCCACGGGAAGGGTACCAAGGCCCTCAGGAGTGGGAGGAGCAC
CGCAACTACGTGAGCAGCTGAAGGGCAGATCGCAAGCTCAAGGAGGAGCTGAGGAGAGGAGTGGCAGCTCAGG
AATGGGCACTACCAAGCAGCTGCTGGCTGGTGGACAGGAGGAGGAGAAGAACCCAGGCCGACCTC
CTGGCCTTCTGCACTCCGAGGTGGACAAGGCAGAGGGTAATGCTGGCTCAAGCTGCCACAGATGTCAGCAGTG
CCTTCTGAGCTTACTCTACAGAAGGTGACCGCTGGAGACTGCCCTACCCCCCAGGGAGAAGCTGTG
AGGAAGGACAAGCGGGATGAGTTGGTGAAGCCATTGAATCAGCCTGGAGACCCCTGAACAATCTGCAGAGAACAGC
CCCAATCACCGCTTACACGGCTCTGATTCATAGAAGGGATCTACCAACAGAAAGGGACAAGGGACATTGTAT
GAGTCACCTTCAAAAGGGGACCAAAACACGAATCAACAGCTCATCTTTCGACCTTCAGCCCCATCATGAAA
GTGAAAATGAAAAGCTCAACATGGCAACAGCTTCAATGTTATCTGTCAGCTCTGACAAAAAGGGTGGACAAGTTC
CGGCACTTCACTGAGAATTTCAGGGAGATGTCATTGAGCAGGATGGAGAGTCCATCTCACTGTGTTACTTGGG
AAAGAAGAAATAATGAAGTCAAAGGAATACTGAAAACACTTCAAAGCTGCCAATTCAAGGAACCTTACCTTCATC
CAGCTGAATGGAGAATTCTCGGGGAAGGGACTTGTGAGTGGAGCCCTCTGGAAGGGAAAGCAACGTCCTTCTC
TTTCTCTGATGTTGGACATCTACTTCACATCTGAATTCTCAATACGTGAGTGGCTGAATACAGGCCAGGGAAAG
GTATTTCAGTCTCTTCACTGAGTCAATCTGGCATAATATACGGCCACCATGATGCACTCCCTCCCTGGAA
CAGCACTGGTCAATAAGAAGGAAACTGGATTGGAGAGACTTGGATTGGGATGACGTGTCAGTATCGGTCAAG
TTCATCAATATAGTGGTTGATCTGGACATCAAAGCTGGGGCGAGAGGATGTGACCTTATCGCAAGTATCTC
CACAGCAACCTCATAGTGTACGGACGCTGTGAGGACTCTTCAACCTCTGGCATGAGAACGGCTGCATGGACAG
CTGACCCCGAGCAGTACAAGATGTGCACTGAGTCCAAGGCATGAGCAGGAGCATCCAGGCCAGCTGGCATGCTG
GTGTTCAAGCAGAGATAGGGCTCACCTCGCAAACAGAAAAGACAAGTAGCAAAAAAAACATGAACTCCAGA
GAAGGATTGTGGAGACACTTTCTTCTGGTCAATTACTGAAAGTGGCTGCAACAGAGAAAAGACTTCCATAAA
GGACGACAAAAGAATTGAGTGGTCAGAGATGAGAAGGCCTCGATTCTCTGTTGGCTTTTACAACAGA
AATCAAATCTCGCTTGCCTGCAAAGTAACTGCAACCCAGTTGACCCCTGTGAAGTGTGACAAAGGCAAGATGCTTGT
AGATTATAAGCTTAATGTTGGAGGTTTGATGGTTTACATACACTGAGACCTGTTGTTGTGCTCATTGA
AATATTCACTGATTAAAGAGCAGTTGAGATCTTCACTGAGATGAGGAGGAGGAGGAGATAGCTTATGCA
GCCTATCAGCAGGGCTCTAGTTCTAGGAATGCTAAATATCAGAAGGGAGGAGGAGGAGATAGCTTATGAT
AGTGAGTACATTAAGTAAAATGGACAGAAAAGAAAAGAACATAATATCGTGTATATTCCCCAAGAT
TAACCAAAAATATCTGTTATCTTGTGCTCTTTAATGCTCTCTTTTCTTTTATTTAAATGCACT
TTTTCTCTGAGTTATGCTGCTTAAATTAACACTTGCACGCCCTAAGAGAGGACACAAGTGGCTAC
ATTTTTATTTTAAAGAGACTTGTGAGATGCTTCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
CCAAGGACATGCCAAATGCTGATTCTGTCAGGCACTGAATGTCAGGCAATTGAGGAGGAGGAGGAGGAG
AATACAGACGTACAGATACTTCTGAGAGTATTCTGAAGAGTATTCTGAAGAGGAGGAGGAGGAGGAG
ACTTTCTGTTACAGAAAAGGAAACTCACTCAGACTGGTGAATCTGAGTGTACCTAAAGTCAGAAACCCACATT
CTCCTCAGAAGTAGGGACCGCTTCTACCTGTTAAATAACCAAGTACCGTGTGAACCAAACAAATCTCTTTC
AAAACAGGGTGCCTCTGGCTCTGGCTTCAAGAAGAAAATGGAGAAAATATATATATATATATATATG
GAAAGATCAATCCATCTGCAAGTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG
TAACTGAATTATTTTAAATAAGCAGTTCTACTCAATCAGGAGGAGGAGGAGGAGGAGGAGGAG
CAAACATTTTTAAATAACAGTTAACATAGAGTGGTTCTCATTCTGAGGAGGAGGAGGAGGAGGAG
ATGCATGAGCTAATTATCTCTTGTGAGTGGCTTCTGCTCACAGTAAACTCATTGTTAAAGCTTCAAGAAC
ATTCAAGCTGTTGGTGTAAATGCAATTGATTGATTGACTGGTAGTTGAAATTAAATTAAACACAGG
CCATGAATGGAAGGTGTTGCAAGCTAATAAAATGATTCTGGATATGAA

FIGURE 72

MMMVRRGLLAWISRVVVLLVLLCAISVLYMLACTPKGDEEQALPRANSPTGKEGYQAVLQEWE
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRPEEKPVRKDKRDELVEAIESALETLNNP
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVKNEKLMAN
TLINVIVPLAKRVDKFQFMQNREMCEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFYPVLFSQY
NPGIIYGHHDADVPPLEQQLVIKKETGFWRDFGFGMTCOYRSDFINIGGFDLDDIKGWGGEDVHLYR
KYLHSNLIVRTPVRGLFHLWHEKRCMDELTPEQYKMCMSKAMNEASHGQLGMLVFRHEIEAHL
RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 73

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGGAGCAAGAGAGTTGTCTGGGATCCA
GAAACCCATGATACCTACTGAACACCGAATCCCTGGAAGGCCACAGAGACAGAGACAGCAAGA
GAAGCAGAGATAAATACACTCACGCCAGGAGCTCGCTCGTCTCTCTCTCTCACTCCTC
CCTCCCTCTCTCTGCCTGTCCTAGTCCTCTAGTCCTCAAATTCCAGTCCCCTGCACCCCTTC
CTGGGACACTATGTTGTTCTCCGCCCTCCTGCTGGAGGTGATTGGATCCTGGCTGCAGATGGG
GTCAACACTGGACGTATGAGGGCCCACATGGTCAGGACCATTGGCCAGCCTCTTACCCGTAGTGT
GGAAACATGCCAGTCGCCATCGATATTCAAGACAGACAGTGTGACATTGACCCGTAGTGT
TGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGCCTTGACCTGCACAACAATGCC
ACACAGTGCAACTCTCTGCCCTCTACCCGTATCTGGTGGACTTCCCCGAAAATATGTAGCT
GCCAGCTCCACCTGCACTGGGTCAAGAAAGGATCCCCAGGGGGTCAGAACACAGATCAACAG
TGAAGCCACATTGAGAGCTCCACATTGTACATTGACTCTGATTCTATGACAGCTTGAGTG
AGGCTGCTGAGAGGCCCTCAGGGCTGGCTGCTGGCATCCTAATTGAGGTGGTGAGACTAAG
AATATAGCTTATGAACACATTCTGAGTCACCTGCATGAAGTCAGGCATAAAGATCAGAACACCTC
AGTGCCTCCCTCAACCTAACAGAGAGCTGCTCCCCAAACAGCTGGGCAGTACTCCGCTACAATG
GCTCGCTCACAACTCCCCCTGCTACCAGAGTGTGCTGGACAGTTTTATAGAAGGTCCCAG
ATTTCAATGGAACAGCTGGAAAAGCTTCAGGGACATTGTTCTCCACAGAACAGGAGCCCTCAA
GCTTCTGGTACAGAACTACCGAGCCCTCAGCCTCTCAATCAGCGCATGGCTTGCTTCTTCA
TCCAAGCAGGATCCTCGTATACCACAGGTGAAATGCTGAGTCTAGGTGAGGAATCTGGTTGGC
TGTCTCTGCCTTCTCTGGCTGTTATTCATTGCTAGAAAGATTCCGAAGAACAGGAGCTGGAAAA
CCGAAAGAGTGTGGTCTTCACCTCAGCACAGCCACGACTGAGGCATAAATTCCCTCTCAGATAC
CATGGATGTGGATGACTTCCCTCATGCCTATCAGGAAGCCTCTAAATGGGTGAGGATCTGG
CCAGAAACACTGTAGGAGTAGTAAGCAGATGTCCTCTCCCTGGACATCTTAGAGAGGAAT
GGACCCAGGCTGTCTTCCAGGAAGAACGTGAGAGCCTCAGCCTCTCCAAACATGTAGGAGGAA
ATGAGGAAATCGCTGTGTTAATGCAGAGANCAAACCTCTGTTAGTGCAGGGGAAGTTGGG
ATATACCCCAAAGTCCTCTACCCCTCACTTTATGGCCCTTCCCTAGATATACTGCAGGATCT
CTCCCTAGGATAAAGAGTTGCTGTTGAAGTTGTATATTTGATCAATATATTGGAAATTAAAG
TTTCTGACTTT

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPASYPECGNNAQSPIDIQTDSVTFPDLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGGLPRKYVAAQLHLHWGQKGSPGGSEHQINSEAT
FAELHIVHYDSDSYDSLSEAAPERQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLENRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCCTGGCGGCCCTGGGACGGGCAGTCCCTGTGTC
TCTGGTGGTTGCCTAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGCCTACA
ATGGACTCCACCAGAGGGTCTCAAGGAGTTAAAGTTACACTGTGCAGTATTTCATCACAA
ATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTCTGTTGCTGACAGCTCC
AGAGAAGTGBAAGAGAAATCCAGAAGACCTCCTGTTCCATGCAACAAATATACTCCAATCTGA
AGTATAACGTGTCTGTGTTGAATACTAAACAGAACGTGGTCCAGTGTGTGACCAACCAC
ACGCTGGTGCACCTGGCTGGAGCCGAACACTCTTACTGCGTACACGTGGAGTCCTCGTCCC
AGGGCCCCCTCGCCGTGCTCAGCCTCTGAGAAGCAGTGTGCCAGGACTTGAAAGATCAATCAT
CAGAGTTCAAGGCTAAATCATCTCTGGTATGTTGCCATATCTATTACCGTGTCTTTT
TCTGTGATGGCTATTCCATCTACCGATATATCCACGTTGCAAAGAGAAACCCAGCAAATT
GATTTGATTATGAAATGAATTGACAAAAGATTCTTGTGCCTGCTGAAAAAAATCGTGA
ACTTTATCACCTCAATATCTGGATGATTCTAAATTCTCATCAGGATATGAGTTACTGG
AAAAGCAGTGTATCCAGCCTAATGATCCTCAGCCCAGCGGGAACCTGAGGGCCCCCTCAGGA
GGAAGAGGGAGGTGAAACATTAGGGTATGCTCGCATTGATGAAATTGGACTCTGAAG
AAAACACGGAAGGTACTTCTCACCCAGCAAGAGTCCCTCAGCAGAACAAATACCCCGGATAAA
ACAGTCATTGAATATGAATATGATGTCAGAACCACTGACATTGTCGGGGCCTGAAGAGCAGGA
GCTCAGTTGCAGGAGGGTGTCCACACAAGGAACATTGGAGTCGCAGGCAGCGTTGGCAG
TCTTGGGCCGAAACGTTACAGTACTCATACACCCCTCAGCTCAAGACTTAGACCCCTGGCG
CAGGAGCACACAGACTCGGAGGAGGGCCGGAGGAAGAGCCATCGACGACCCTGGTCAGTGG
TCCCCAAACTGGCAGGCTGTGTATTCTTCGCTGTCCAGCTCGACCAGGATTCA
AGAGGGCTGCG
AGCCTCTGAGGGGGATGGCTGGAGAGGGAGGTCTTCTATCTAGACTCTATGAGGAGCCGGCT
CCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCTATGGAGGAATGGGGTTATA
TGTGCAGATGGAAAACTGATGCAACACTTCTTTGCCTTGTGCAAACAAAGTGAG
TCACCCCTTGATCCCAGCCATAAAAGTACCTGGGATGAAAGAAGTTTCCAGTTGTCA
CTGTGAGAATTACTTATTCTTCTATTCTCATAGCACGTGTGATTGGTCATGCATGTA
GGTCTCTAACATGATGGTGGCCTCTGGAGTCCAGGGGCTGGCCGGTTGTTCTATGCAGAGAA
AGCAGTCAATAATGTTGCCAGACTGGGTGCAGAATTATTCAAGGTGGGTGT

FIGURE 76

MSYNGLHQRVFKELKLLTLCISISSQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY
SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNLTYCVHVESFVPGPPRRAQPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNEDKRFVPAEK
IVINFITLNISDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEVKHLGYASHLMEIFC
DSEENTEGTSLTQQESLSRTIPPDKTVIEYEDVRTTDICAGPEEQELSLOEEVSTQGTLLESQA
ALAVLGPTLQYSYTPQLQDPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSFDQDS
EGCEPSEGDGLEEGLLSRLYEEPAPDRPPGENETYLMQFMEEWGLYVQMen

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCAGGCCGAGGACTCCAGCGTGCCTAGGTCTGGCATCCTGCACTTGCTGCCCTTGACAC
CTGGGAAGATGGCCGGCCGTGGACCTTCACCCCTCTGTGGTTGCTGGCAGCCACCTTGATC
CAAGCCACCCCTCAGTCCCAC TGCA GTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC
ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGTGCTCAGTGCCATGC
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGCAGCCTGGTGAACACCGTCCTGAACCACATC
ATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAAGGCCCTGGCCAATGACCA
GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
TCGTGGAGTTCCACATGACGACTGAGGCCAAGCCACCATCCGCATGGAACCAGTGCAGTGGC
CCCACCCGCCTGGTCCTCAGTGA CTGTGCCACCAGCCATGGGAGCCTGCGCATCCA ACTGCTGTA
TAAGCTCTCCTTCCTGGTGAACGCCCTAGCTAACGAGGTCAAGCTCC TAGTGCCATCCCTGC
CCAATCTAGTGAAAACCAGCTGTGTCCGTGATCGAGGCTTCCTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCATTTCCTCAGCATTGACCGTCTGGAGTTGACCTTGTGA
TCCTGCCATCAAGGGTACACCAATTCAAGCTCTACCTGGGGCCAAGTTGTTGGACTCACAGGGAA
AGGTGACCAAGTGGTCAATAACTCTGCAGCTCCCTGACAATGCCACCCCTGGACAACATCCCG
TTCAGCCTCATCGTAGTCAGGACGTGGTGAAGAGCTGCAGTGGCTGCTGTGCTCTCCAGAAGA
ATTCAATGGTCCCTGGACTCTGTGCTTCCTGAGAGTGCCATGGCTGAAGTCAAGCATCGGGC
TGATCAATGAAAAGGCTGCAGATAAGCTGGATCTACCCAGATCGTAAGATCTTAACACTCAGGAC
ACTCCCAGTTTTATAGACCAAGGCCATGCCAAGGTGGCCAACTGATCGTAGCTGGAGTGT
TCCCTCCAGTGAAGGCCCTGCCCTTGTACCCCTGGCATGAGCCAGCTCGGAAGCTCAGT
TTTACACCAAAGGTGACCAACTTAACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTG
ATGAACCTGGGATTGGCTGGTCCAACCTGATGTTGCTGAAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCCGAACCAGAAATGGCAAATTAAAGATCTGGGTCCCAGTGTCAATTGGTGAAGG
CCTTGGGATTGAGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTGTGCTTACTCCAGCCTCC
TTGTGGAAAACCCAGCTCCTGTCTCCAGTTGAAGACTGGATGGCAGCCATCAGGGAAAGGCTGG
GTCCCAGCTGGAGTATGGGTGTGAGCTATAGACCATCCCTCTGCAATCAATAAACACTTG
CCTGTGAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTOELKDHNATSILQQPLLSAMREK
PAGGIPLVLSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTTEAQATIRMDTSASGPTRLVSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQQLYLGAKLLDSQGKVT
KWFNNSAASLTMPTLDNIPFSLIVSQDVVKAAVAAVLSPEEFMVLLDSVLPEAHRLKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSEALRPLFTLGIEASSEAQFYT
KGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEIHSILLPNQNGKLRSGVPVSLVKALG
FEAAESSLTKDALVLTPASLWKPPSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTACTGAGAGGTCTGCCATGGCCTCT
CTTGGCCTCCAACTTGTGGCTACATCCTAGGCCTCTGGGCTTTGGGCACACTGGGTGCCAT
GCTGCTCCCCAGCTGGAAAACAAGTTCTATGTCGGTGCCAGCATGTGACAGCAGTTGGCTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC
ACCCTCTGGCCTGCCGCTGACATCCAGGCTGCCAGGCCATGATGGTGACATCCAGTGCAAT
CTCCTCCCTGGCCTGCATTATCTCTGTGGGGCATGAGATGCACAGTCTCTGCCAGGAATCCC
GAGCAAAGACAGACTGGCGGTAGCAGGTGGAGTCTTTCATCCTGGAGGCCTCTGGGATTC
ATTCCCTGTTGCCTGGAATCTCATGGGATCCTACGGACTTCTACTCACCCTGGTGCTGACAG
CATGAAATTGAGATTGGAGAGGCTCTTACTTGGGATTATTCTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTCTGCTCATCCAGAGAAATCGCTCAAACACTACGATGCCCTAC
CAAGCCCAACCTCTGCCACAAGGAGCTCCAAGGCCCTGGTCAACCTCCAAAGTCAAGAGTGA
GTTCAATTCCCTACAGCCTGACAGGGTATGTGTGAAGAACCAAGGGCCAGAGCTGGGGGTGGCTG
GGTCTGTAAAAACAGTGGACAGCACCCCGAGGCCACAGGTGAGGGACACTACCACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTGGCATTGGATTGAGCAAAGGCAGAAATGGGG
GCTAGTGTAAACAGCATGCAGGTTGAATTGCCAAGGATGCTGCCATGCCAGCCTTCTGTTTCC
TCACCTGCTGCTCCCTGCCCTAAGTCCCCAACCTCAACTGAAACCCATTCCCTTAAGCCA
GGACTCAGAGGATCCCTTGCCCTCTGGTTACCTGGACTCCATCCCCAAACCCACTAATCACA
TCCCACGTGACTGACCCCTCTGTGATCAAAGACCCCTCTCTGGCTGAGGTTGGCTTAGCTCATT
GCTGGGATGGGAAGGAGAAGCAGTGGCTTTGTGGCATTGCTCTAACCTACTCTCAAGCTTC
CCTCCAAAGAAACTGATTGGCCCTGGAACCTCCATCCACTCTGTTATGACTCCACAGTGTCCA
GACTAATTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGAAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGACATTAAAAAAATA

FIGURE 8o

MASLGLQLVGYILGLLGLLGTIVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMVTSsAISSLACII SVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAWNLHGILRDFYSPLVPSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCACCGTCCGCCTCTCCCTCTGCTGGACCTCCTCGTCTCCATCTCTCCCTCTTC
CCCGCGTTCTCTTCCACCTTCTCTTCCACCTAGACCTCCCTCGCCCTCCCT
GCCAACCGCTGCTCCCTGGCCCTCTCCGACCCGCTAGCAGCAGACCTCCTGGGGTCTGTGG
GTTGATCTGTGGCCCCGTGCTCCGTGCTCCCTCGTCTCCCTCCGACTCCGCTCCGG
ACCAGC GCCCTGACCTGGGAAAGGAGGGATGGTCCCGAGGTGAGGGTCCCTCCCTGCTGGGA
CTCGCGCTGCTCTGGTCCCCCTGGACTCCCACGCTCGAGCCCCGCCCAGACATGTTCTGCCTT
CCATGGGAAAGAGATACTCCCCCGCGAGAGCTGGCACCCCTACTTGAGGCCACAAGGCTGATGT
ACTGCCCTGCCGTGTACCTGCTCAGAGGGGCCCATGTGAGTTACCGCCTCCACTGTCCGCC
GTCCACTGCCCTGACGGAGCCACAGCAATGCTGTCCAAGTGTGTGGAACCTCACAC
TCCCTCTGGACTCCGGGCCCCACCAAAGTCTGCCAGCACAAACGGGACCATGTACCAACACGGAG
AGATCTTCAGTGCCCATGAGCTGTTCCCTCCGCCCTGCCAACAGTGTGTCCCTGCAGCTGC
ACAGAGGGCCAGATCTACTGCCCTCACAACTGCCCGAACCGAGGCTGCCAGCACCCCTCCC
ACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGAAGTGAAGCAATCGGATGAAGAGGACA
GTGTGCAGTCGCTCCATGGGTGAGACATCCTCAGGATCCATGTTCCAGTGTGCTGGGAGAAAG
AGAGGCCGGGACCCCCAGCCCCACTGCCCTAGGCCCTAGGCCCTCTGAGCTTCATCCCTGCCACTT
CAGACCCAAGGGAGCAGGCAGCACAACTGTCAGATCGCTGAAGGAGAAACATAAGAAAGCCT
GTGTGCATGGGGGGAAAGACGTACTCCACGGGGAGGTGTGGCACCCGCCCTCCGTGCCCTCGGC
CCCTGCCCTGCATCTATGCCACTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
CACCGAGTACCCCTGCCGTACCCGAGAAAGTGGCTGGGAAGTGTGCAAGATTGCCAGAGG
ACAAAGCAGACCCCTGCCACAGTGAGATCAGTTCTACCAAGGTGTCCAAGGCACCGGGCCGGTC
CTCGTCCACACATCGGTATCCCCAAGCCAGACAACCTGCGTCGTTGCCCTGGAACACGAGGC
CTCGGACTTGGTGGAGATCTACCTCTGGAAAGCTGGTAAAAGATGAGGAAACTGAGGCTCAGAGAG
GTGAAGTACCTGGCCAAGGCCACACGCCAGAACATCTCCACTTGACTCAGATCAAGAAAGTCAG
GAAGCAAGACTTCCAGAAAGAGGCACAGCAGTCCGACTGCTCGTGGCCCCCACGAAGGTCACT
GGAACGTCTCCTAGCCCAGACCCCTGGAGCTGAAGGTACGCCAGTCCAGACAAAGTGAACCAAG
ACATAACAAAGACCTAACAGTTGCAGATATGAGCTGTATAATTGTTATTATATATAATAAAA
TAAGAAGTTGCATTACCCCTCAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKCVEHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELP
PSRLPNQCVLCSCTEGQIYCGLTCPEPGCPAPIPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPGBTAPTGLSAPLSFIPRHFRPKAGSTTVKIVLKEKKKACVHGGKTY
HGEVWHPAFRAFGPLCILCTCEDGRQDCQRVTCPTEYPCRHPEKVAGKCKKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLSDQESQEARNPERGTALPTARWPRRSLERLSPDPGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTGTCGATGGAGTAGACTCTCAGAACAGCGCAGTTGCCCTCGCTCACGAGAGCCTCTCC
GTGGCTTCCGCACCTTGAGCATTAAGGCCAGTTCTCCTCTCTAATCCATCCGTACCTCTCCTGTCA
TCCGTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCCTCATGCTCAGTTGGTCTGAGTC
TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTGGCCAGACAAGCCTGCCAGGCCTGGTGGGGAG
GACGCAGCATTCTCCTGTTCTCTCTAAGACCAATGCAGAGGCCATGGAAGTGCAGGTTCTCAGGGG
CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTATGCAGATGCCACAGTATC
AAGGCAGGACAAAACGGTGAAGGATTCTATTGCGGAGGGCGCATCTCTGAGGCTGGAAAACATTACT
GTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCCAGTCTTACTACCAGAAGGCCATCTGGGAGCT
ACAGGTGTCAACTGGCTCAGTTCTCTCATTTCCATACGGGATATGTTGATAGAGACATCCAGCTAC
TCTGTCAGTCCTCGGGCTGGTCCCCGGCCACAGCGAAGTGGAAAGGTCCAAGGACAGGATTGTCC
ACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTGATGTGGAGATCTCTGACCGTCCAAGAGAA
CGCCGGAGCATATCCTGTTCCATGCCATGCTCATCTGACCGAGAGGTGGAATCCAGGGTACAGATAG
GAGATACTTTTCGAGCCTATATCGTGGCACCTGGCTACAAAGTACTGGAATACTCTGCTGTGGCCTA
TTTTTGGCATTGTTGGACTGAAGATTCTCTCCAAATTCCAGTGGAAATCCAGGCCAACTGGACT
GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAG
AGACGGCTCACCCGAAGCTCTGCGTTCTGATCTGAAAACGTAAACCCATAGAAAGCTCCCAGGAGGTG
CCTCACTCTGAGAAGAGATTACAAGGAAGAGTGTGGTGGCTCTCAGAGTTCCAAGCAGGGAAACATTA
CTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGGTGGAGTGTGCCCCGATGATGTGGACAGGAGGA
AGGAGTACGTGACTTGTCTCCGATCATGGTACTGGGTCTCAGACTGAATGGAGAACATTGTATTTC
ACATTAATCCCGTTTATCAGCGTCTCCCAAGGACCCACCTACAAAAAATAGGGTCTTGGACT
TGAGTGTGGACCATCTCCTCTCAACATAATGACCAACTCCCTTATTATACCCGACATGTCGGTTG
AAGGCTTATTGAGGCCCTACATTGAGTATCGCTCTATAATGAGCAAATGGAACCTCCATAGTCATCTGC
CCAGTCACCCAGGAATCAGAGAAAGAGGCCTCTGGCAAAGGGCCTGCAATCCAGAGACAAGCAACAG
TGAGTCCTCCTCACAGGCAACCACGCCCTCCCTCCCAGGGTGAATGTAGGATGAATCACATCCACAT
TCTCTTTAGGGATATAAGGTCTCTCTCCAGATCAAAGTCCCGCAGCAGCCGCCAAGGTGGCTTCCA
GATGAAGGGGACTGCCGTCCACATGGAGTCAGGTGTATGGCTGCCCTGAGCTGGAGGGAAAGAAG
CTGACATTACATTAGTTGCTCTCACTCCATGGCTAAGTGTCTTGAATACCAACCTCTCAGGTGAAG
AACCGTCAGGAATTCCCATCTCACAGGCTGTGGTAGATTAAGTAGACAAGGAATGTGAATAATGCTTAG
ATCTTATTGATGACAGAGTGTATCCTAATGGTTGTTCAATTATATTACACTTCAGTAAAAAA

FIGURE 84

MALMLSIVSLLKLGSGQWQVFGPDKPVQALVGEDAASFCLSPKTNAEAMEVRFFRGQFSSVVH
LYRDGKDQPFMOMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYYQKAIWELQ
VSALGSVPPLISITGYVDRDIQLLCQSSGWFPRTAKWKGPQGQDLSTDRTNRDMHGLFDVEISL
TVQENAGSISCSMRRAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAELDWRRKHGQAEELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKPQEVPHSEKRF
TRKSVVASQSFOAGKHYWEVDGGHNRWRVGVCRDDVDRRKEYVTLSPDHGYWVRLNGEHLYFT
LNPRFISVFPPPTKIGVFLDYECGTISFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
PIVICPVTQESEKEASWQRASAIPETSNSESSSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGGGCCCTGGCACCTCAACCCAGACATGCTGCTGCTGCTGCCCT
GCTCTGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTTCCGTGA
CGGTGCAGGAAGGCCTGTGTCCATGTGCCCTGCTCCTCCTACCCCTCGATGGCTGGATT
TACCCCTGGCCCAGTAGTCATGGCTACTGGTCCGGGAAGGGGCCAATACAGACCAGGATGCTCC
AGTGGCCACAAACAACCCAGCTGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTG
GGGACCCACATACCAAGAATTGCACCCCTGAGCATCAGAGATGCCAGAAGAAGTGTGCGGGGAGA
TACTTCTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATACCGGCTCTGTGAA
TGTGACAGCCTGACCCACAGGCCAACATCCTCATCCAGGCACCCCTGGAGTCCGGCTGCC
AGAATCTGACCTGCTCTGTGCCCTGGCCTGTGAGCAGGGACACCCCTATGATCTCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCACCACCCGCTCCTCGGTGCTCACCTCATCCCACA
GCCCCAGGACCATGGCACCAGCCTCACCTGTCAGGTGACCTTCCCTGGGCCAGCGTGACCACGA
ACAAGACCGTCCATCTCAACGTGCTCACCGCCTCAGAACCTGACCATGACTGTCTCCAAGGA
GACGGCACAGTATCCACAGTCTGGAAATGGCTCATCTGTCACTCCCAGAGGGCAGTCTCT
GCGCCTGGTCTGTGAGTTGATGCAGTGCAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCCCTGACCCCTGTGCCCTCACAGCCCTAAACCCGGGGTGCTGGAGCTGCCTGGTGAC
CTGAGGGATGCAGCTGAATTACCTGCAGAGCTCAGAACCTCTCGCTCTCAGCAGGTCTACCT
GAACGTCCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTCGGGGAGCTGGAG
CCACAGCCCTGGTCTTCTGTGCTCATCTCGTTGAGGTCTGCAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGCTGGGAGATACGGGCTAGAGGATGCAAACGCTGTCAAGGGTTC
AGCCTCTCAGGGCCCTGACTGAACCTGGCAGAAGACAGTCCCCAGACCAGCCTCCCCAG
CTTCTGCCGCTCTCAGTGGGAAAGGAGAGCTCAGTATGCATCCCTCAGCTCCAGATGGT
AAGCCTGGACTCGCGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAGAAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCTCCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTGTAGAATTAAACAGCCCTAACGTGATGAGCTATGATAACACTATGAATTATG
TGCAGAGTAAAAGCACACAGGCTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCC
TCCCTTTATTTTTAACTAAAGACAGACAAATTCTA

FIGURE 86

MLLLLLPLLWGRERERAEGQTSKLLTMQSSVTQEGLCVHPCFSYPSHGWIYPGPVVHGYWFREG
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLISRARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCSVPWACEQGTPPMISWIGTSVSPLDPSTTRS
SVLTLIPOQPQDHGTSLTCQVTFFPGASVTTNKTVHLNVSYPPQNLTMTVFQGDGTVSTVLGNSSL
SLPEGQSLRLVCAVDADSNPPARLSLSWRGLTLCPSQPSNPVGYLELPWVHLRDAEFTCRAQNP
LGSQQVYLNVSLSQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGDTGIE
DANAVRGASASQGPLTEPWAEDSPPDQPPPASARSSVGEHELQYASLSFQMVKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGTGAAGGAGCTCTGTAC
CCAAGGAAAGTCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTCCTGCTGTTTC
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTGCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGTGCATTGA
TGGCCTGTATTTCTCCGCACTGAGAATGGTGTATCTACCAAGACCTCTGTGACATGACCTCTG
GGGGTGGCGGCTGGACCCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGAGTGCACGGTG
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGACGGCACTGGC
CAACTACAACACCTTGGATCTGCAGAGGGCCACGAGCGATGACTACAAGAACCCCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGCATCTGGCACGTGCCAATAAGTCCCCATGCAGCACTGG
AGAAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTCCTCAGACACTGGGACATAATCT
GTTTGGCATCTACCAAGAAATATCCAGTGAAATATGGAGAAGGAAAGTGTGACTGACAACGCC
CGGTGATCCCTGTGGTCTATGATTTGGCGACGCCAGAAAACAGCATCTTATTACTCACCTAT
GGCCAGCGGAATTCACTGCGGATTTGTTAGTTCAAGGGTATTTAATAACGAGAGAGCAGCCAA
CGCCTTGTGCTGGAATGAGGGTACCGGATGTAACACTGAGCATCACTGCATTGGGGAGGAG
GATACTTCCAGAGGCCAGTCCCAGCAGTGTGGAGATTTCTGGTTTGATTGGAGTGGATAT
GGAACTCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TTGAGAGTTTGTGGAGGGAACCCAGACCTCTCCCAACCATGAGATCCCAAGGATGGAGAA
CAACTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWLTVASVHENDMRGKCTVGDRSSQQGSKADYPFGDGNWANYNTFGSAEAAT
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTDNGPVIKVYDFGDAQKTASYSPYQREFTAGFVQFRVFNNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTGTCGGCTGCGGGGAGACTCAGGAGTCGCTGTCTGAACCTCCAGCCTCAGAGAC
CGCCGCCCTGTCCCCGAGGGCATGGCCGGTCTCAGGGCTGTGCCCTCTCGCTTCCTGACG
CTCCTGGCGCATCTGGTGGTCGTACACCTATTCTGGTCCCAGGACAGAACATAACAGGCCTG
CCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTCAGCTGGTGGCCGCGCTT
CTGTCACCCCTGGGCCTTTGCAGTGGAGCTGGCCGGTTCCCTCTCAGGAGTCTCCATGTTAAC
AGCACCCAGAGCCTCATCTCCATTGGGCTCACTGTAGTGCATCCGTGGCCCTGTCCTTCTTCAT
ATTCGAGCGTGGAGTGCACTACGTATTGGTACATTTTGTCCTCTGCAGTGCCCTTCAGCTG
TCACTGAAATGGCTTATTCGTCACCGTCTTGGCGTGAAAAAGAAACCCTTCTGATTACCTTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGCAAGGGCCGCTCGTATTCCCTGGAAGAAGGAAG
GCATAGGGCTCGGTTTCCCTCGGAAACTGCTCTGCTGGAGGATATGTGTTGAAATAATTACG
TCTTGAGTCTGGATTATCCGCATTGTATTAGTGCCTTGTAATAAAATGTTTAGTAACA
TTAAGACTTATATACAGTTAGGGACAATTAAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVTEMALFV
TVFGLKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCGAAAAGAGAAGGGGAGAGCGAGGGACGAGAGCGGAGGAAGATGCAACTGAC
TCGCTGCTGTTCTGGTGCAGGGTAGCCTCATCTGGTATCTGTGGCCAGGATGATG
GTCCTCCCGCTCAGAGGACCTGAGCGTATGACCACGAGGGCAGCCCCGGCCCAGGTCCT
CGGAAGCGGGGCCACATCTCACCTAACGTTCCGGCCATGGCAATTCCACTCTCCTAGGGCTGCT
GGCCCCGCTGGGGAGGCTTGGGCATTCTGGCAGCCCCCAACCGCCGAACCACAGCCCC
CACCTCAGCCAAGGTGAAGAAATCTTGCTGGGCAGTCTACTCCAACATCAAGACGGTG
GCCCTGAACCTGCTCGTCACAGGGAAAGATTGTGGACATGGCAATGGACCTTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAAACATCTCATCAGCCTCGTGCCTCCAGTAAAGCTGTAG
AGTTCCACCAGGAACAGCAGATCTCATCGAAGCCAAGGCCCTCAAATCTCAACTGCCGGATG
GAGTGGGAGAAGGTAGAACGGGCCGGACCTCGTTGCACCCAGGCCAAGATCTG
CTCCCGAGACCACGCTCAGAGCTCAGGCCACTGGAGCTGCTCCCAGCCCTCAAAGTCGTCTGTG
TCTACATGCCCTCTACAGCAGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAACATAC
CATAGTGAACCCCTACTACCCATCTGGGTGACCCGGGGCAGGCCACAGAGGCCAGGCCAGGGC
TGGAAAGGACAGGCCCTGCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAAGGGTTGGCCTC
AGGCAGGGAGGGGGTGGAGACGAGGAGATGCCAAGTGGGCCAGGCCAAGTCTCAAGTGGCAG
AGAAAGGGTCCAAGTGTGGTCCAAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG
AGGAGGGAGTGGCTCTGTGCAGCCTCACAGGGCTTGCCACGGAGCCACAGAGAGATGCTGG
TCCCCGAGGCCCTGGGCAGGCCATCAGTGTGGCCCCAGATCAAGTCATGGAGGAAGCTAAC
CCTGGTTCTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTCACTCAGTGTGGACA
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCCCTAGGAGCCAGTCAGCAGGGTGGGGTGG
GCCAGAGGGAGCTCCAGCCCTGCCTAGTGGCGCCCTGAGCCCTGTGCTGTGAGCATGG
CATGAGGCTGAAGTGGCAACCCCTGGGTCTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC
CAGGCCACCCCTTCAAATCCCTTCTGCCAGTACTCCCCCTGTACCACCCATTGCTGATG
GCACACCCATCCTTAAGCTAACGACAGGACGATTGTGGTCTCCCACACTAACGGCCACAGCCCAC
CGCGTGTGTGTGTCCTCTCCACCCCAACCCCTGCTGGCTCCTCTGGAGCATCCATGCCCG
GAGAGGGTCCCTAACAGTCAGCCTCACCTGTCAGACCGGGGTTCTCCGGATCTGGATGGCG
CGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGCCAGAGCATGTGCTGGATCTGTC
TGTGTGCTGTCTGTGGGTGGGGAGGGGAGGGAAAGTCTGTGAAACCGCTGATTGCTGACTTT
TGTGTGAAGAATCGTGTCTTGGAGCAGGAATAAGCTTGGGGGGCA

FIGURE 92

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNNTF
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMEWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPFKVVVCYIAFYSTDYRLVQKVCPDNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCCATGACTGCGGCCGTGTTCTCGGCTGCGCCTCATTCGCCTCGGGCTGCGCTGCC
CTTTATGTCTTCACCATGCCATCGAGCCGTTGCGTATCATCTTCCTCATGCCGGAGCTTCTT
CTGGTTGGTGTCTACTGATTCGTCCCTGTTGGTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTGAGCGTTGTCATGTCTATCCAAGAA
ATGTTCCGATTTGCATATTATAAACTCTTAAAAAAAGCCAGTGAAGGTTGAAGAGTATAAACCC
AGGTGAGACAGCACCCCTATGCGACTGCTGGCCTATGTTCTGGCTGGGCTTGGAAATCATGA
GTGGAGTATTTCTTGTGAATACCCATCTGACTCCTGGGCCAGGCACAGTGGGATTCA
GGAGATTCTCCTCAATTCTCCTTATTCAAGCTTCATGACGCTGGTCATTATCTGCTGCATGT
ATTCTGGGCATTGTATTTTGATGGCTGTGAGAAGAAAAAGTGGGCATCCTCCTTATCGTTC
TCCTGACCCACCTGCTGGTGTCAAGCCCAGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCG
TCAGCATTATAATCCTGGTGCTCATGGCACCTGGCATTCTTAGCTGCGGGAGGCAGCTGCCG
AAGCCTGAAACTCTGCCTGCTCTGCCAAGACAAGAACTTCTTACAACCAGCGCTCCAGAT
AACCTCAGGAACCAGCACTTCCAAACCGCAGACTACATCTTAGAGGAAGCACAACTGTGCCT
TTTCTGAAAATCCCTTTCTGGTGAATTGAGAAAGAAATAAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAGPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFRFAYYKILLKKASEGLKSINPGETAPSMRLLAYVSGLGFMSGV
FSFVNTLSDSLGP GTVGIHGDSPQFFLYSAFMTLVIILLHVFWGIVFFDGCEKKKGILLIVLLT
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSLKLCLLCQDKNFLYNQRSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTCACCAGAGTAAACTGAGAAACCAACTGGACCTTGAGTATTGTACATTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTGTCTCTAGGATCAAC
TCGGTCATTACCACAGCTCAAACCTGCTTGGACTCCCTCCCACAAAACTGGCTCCGGATCAGG
GAACACTACCAACCAACAGCAGTCAAATCAGGTCTTCCTTAAAGTCTGATAACCATTAACA
CAGATGCTCACACTGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTGTCACACAACTTGGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCCATTCTGTTCCGGAGGCATCCTGCCACCAGTCAGGCAGGGC
TAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAGCAGGTGTAATCCTGCCACCCAGG
GAACCCAGCAGGCCCTCCCAACTCCAGTGGCACAGATGACGACTTGCAGTGACCACCC
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA
GTAAGCTGTTCAAATTTTCAACTAAGCTGCCTCGAATTGGTGATACATGTGAATCTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAAATTCTTAATT
TACCTGAAATATTCTTGAAATTCAAGAAATATGTTCTATGTAGAGAAATCCAACTTTAAAAAA
CAATAATTCAATGGATAAAATCTGTTGAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAAACATATTGGAAAACTGGAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNOVFPSSLIPLTQM
LTLPDPDLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNATQGTPAGRLPTPSG
TDDDFAVTTPAGIQQRSTHAIIEEATTESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

GCTCAAGTGCCTGCCCTGCCACCCAGCCCAGCCTGCCAGAGCCCCCTGGAGAAGGAGCTCT
CTTCTGCTGGCAGCTGGACCAAGGGAGCCAGTCCTGGCGCTGGAGGGCTGTCCGACCATG
GTCCCTGCCCTGGCTGGCTGTTGTCTCGTCCCCAGGCTCTCCCAAGGCCAGCTGC
AGAGCTGTCTGTGGAAGTCCAGAAAACATATGGTGGAAATTCCCTTATAACCTGACCAAGTTGC
CGCTGCCCGTGGAGGGGCTGAAGGCCAGATCGCTGTCAGGGACTCAGGCAAGGCAACTGAG
GGCCCATTTGCTATGGATCCAGATTCTGCTCCCTGTCAGGGACTCAGGCAAGGCAACTGAG
GCAGGCAGAGTACCAAGCTACAGGTACCCCTGGAGATGCAGGATGGACATGTCCTGGGGTCCAC
AGCCTGTGCTTGCACTGTGAAGGATGAGAAATGACCAAGGTGCCCCATTCCTCTGAGGCTTCAGACCG
AGAGCTGGCTGAGCCGGGGTACCCAGGCCCTGGCAGTCCCCCTCTCCTGAGGCTTCAGACCG
GGATGAGGCAGGCCACGCCAACCTGGATCTCGATTCCACATCTGAGCCAGGCTCCAGGCCAGC
CTTCCCCAGACATGTCAGCTGGAGCCCTGGCTGGGGCTCTGGCCCTCAGCCCCAAGGGGAGC
ACCAGCCTTGACCAACGCCCTGGAGAGGACCTACAGCTGGTACAGGTCAAGGACATGGGTGA
CCAGGCCCTCAGGCCACCCAGGCCCTGGAGAGGACCTACAGCTGGTACAGGTCAAGGACATGGGTGA
CCCTAGAGCCTATCCACCTGGAGAGAAATCTCAAAGTCTATACCCGACCAATGGCCAGGTA
CACTGGAGTGGGGGTATGTGCACTATCACCTGGAGAGCCATCCCCGGGACCTTGAAGTGAA
TGCAGAGGGAAACCTCTAGTGAACAGAGAGCTGGACAGAGAAGGCCAGGCTGAGTACCTGCTCC
AGGTGCGGGCTCAGAAATCCATGGCAGGACTATGCGCCCTCTGGAGCTGACGTGCTGGTGA
ATGGATGAGAAATGACAACGTGCTATCTGCCCTCCCCGTGACCCACAGTCAGCATCCGTAGCT
CAGTCCACCAAGGTACTGAAGTGAAGTAGACTGTCAGCAGAGGATGCAGATGCCCGGCTCCCCA
ATTCCCACGTTGTATCAGCTCTGAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCTTC
CAGGTGGACCCCACCTCAGGCACTGTGACGCTGGGGTGTCCACTCCGAGCAGGCCAGAACAT
CTGCTCTGGTGTGGCATGGACCTGGCAGGCCAGAGGGTGGCTTCAGCAGCACGTGTGAAG
TCGAAGTCGAGTCACAGATATACTGATCACGCCCTGAGTTCATCACTTCCAGATTGGGCT
ATAAGCCTCCCTGAGGATGTGGAGCCGGACTCTGGTGGCATGCTAACAGCATTGATGCTGA
CCTCGAGCCCGCCTCCGCTCATGGATTGGCATTGAGAGGGGAGACACAGAAGGGACTTTG
GCCTGGATTGGGAGCCAGACTCTGGGATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG
GCAGCTCAAGTCATGAGGTGGTGGTGGCAGAGTGTGGCAGAGCTGGTGGGGCCAGGCC
AGGCCCTGGAGCCACGCCACGGTACTGTGCTAGTGGAGAGAGTGTGCCACCCCCAAGTTGG
ACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTGCCAGGCCCTTCTGCTGACCATC
CAGCCCTCCGACCCATCAGCCGAACCCCTCAGGTTCTCCTAGTCATGACTCAGAGGGCTGGCT
CTGCACTGGAGAAATTCTCCGGGAGGTGCAACCCGCCCAGTCCCTGCAAGGGGCCAGCCTGGG
ACACCTACACGGTGTGGAGGCCAGGATACTGCCCTGACTCTGCCCTGTGCCCTCCAA
TACCTCTGACACCCGCCAAGACCATGGCTGACTCGTGGACCTGCAAGGGACCCGATCT
GGCAGTGGGACGGTCCCTACAGCTCACCCCTGGTCCCAACCCACGGTGCACGGGATTGGC
GCCTCCAGACTCTCAATGGTCCCATGCCCTACCTCACCTGGCCCTGCATTGGGTGGAGGCCACGT
GAACACATAATCCCGTGGTGGCAGCCACAATGCCAGATGTGGCAGCTCTGGTTCAGTGT
CGTGTGCGCTGCAACGTGGAGGGCAGTGCATGCCAGGGTGGCCATGAAGGGCATGCCA
CGAAGCTGTGGCAGTGGCATCTGTAGGCAACCTGGTAGCAATAGGAATCTCCATCCTC
ATTTTACCCACTGGACCATGTCAGGAAGGAAGGACCCGATCAACCCAGCAGACAGCTGCCCT
GAAGGGACTGTCTGAATGGCCCAAGGCAGCTAGCTAGCTGGGAGCTGGCCTCTGGCTCCATCTGAG
TCCCCCTGGGAGAGAGGCCAGCAGCAAGATCCAGCAGGGGACAGGACAGAGTAGAAGCCCTCCA
TCTGCCCTGGGGTGGAGGCCACCATCACCACATCACCAGGCACTGTGCAAGAGCCTGGACACCAACTT
TATGGACTGCCCATGGAGTGCTCCAAATGTCAGGGTGTGCCCCATAATAAGCCCCAGAGAA
CTGGGCTGGGCCATGGAAAAAAGAAAAAAAAAAGAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSVPQALPKAQPAELSVEVPENYGGNFLYLTKLPLPREGAEGQIVLSGDSKGAT
EGPFAMDPDSGFLLVTRALDREEQAELYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQPSPDMFQLEPRLGALALSPKG
STS LDHALERTYQLLVQVKDMGDQASGHQATATVEVSIIESTWVSLPIHLAENLKVLYPHHMAQ
VHWSGGDVHYHLESHPGPFEVNAEGNLYVTRELDREAQEYLLQVRAQNSHGEDYAAPPLEHVL
VMDENDNVPICPPRDPPTVSIPELSPPGTEVTRLSAEDADAPGSNSHVYYQLLSPEPEDGVGRA
FQVDPTGSVTLGVPLRAGQNILLLVLAMDLAGAEGGFSSTCEVEAVTDINDHAPEFITSQIG
PISLPEDVEPGTLVAMLTAIDADLEPAFRIMDFAIERGDTEGTFGLDWEPDSGHVRLRLCKNLSY
EAAPSHEVVVVVQSVAKLVGPGPGATATVTVLVERVMPPPQLDQESYEASVPISAPAGSFLLT
IQPSDPISRTLRFSLVNDSEGWLCLIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALTLPVPS
QYLCTPRQDHGLIVSGPSKDPDIASGHGPYSFTLGPNPQVQRDWRLQTLNGSHAYLTLALHWEP
REHIIPVVVSHNAQMWMQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGTVAIGIFLI
LIFTHWTMSRKKDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCACGCCGT
AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCTTTA
TCTCTCACCTCAAGTCCCCCTTCTCAAGAATCCTCTGTTCTTCAGCCCTAAAGTCTGGTAC
ATCTAGGACCCAGGCATCTGCTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAGGAATG
TTCTCCTATGTTGGTACTATTGCAATTAGAAGCTGCAACAAATTCCAATGAGACTGACCC
TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGTCCAG
TGTGACCTCCAGTGGGTCAAGCACAGCCACCATCTCAGGGTCCAGCGTGAACCTCAAATGGGTCA
GCATAGTCACCAACTCTGAGTTCCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACACCTCCAGTGG
GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCCTCAGTGGGCCAGCACAGTCACCAACT
CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC
AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGCTAGCACAGCCAC
CAACTCTGACTCCAGCACAACTCCAGTGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGCCAGCACAGCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT
GCCACCAACTCTGAGTCCAGCACACCCCTCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCAAATGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACAGCCACCAACTCTGAGTCCAG
GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGCCAGCACTGCCACCAACTCTGAG
TCCAGCACGACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGCACCTCCAGTGG
GGCTAGCACAGCCACCAACTCTGACTCCAGCACAGCACACCCCTCAGTGGGCCGGCACAGCCACCAACT
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCACTGAGTCCAGCACAGCACCCCTCC
AGTGGGCCACACAGCCACCAACTCTGAGTCCAGTGGGCCAGCAACTCTGAGTCCAGCACACAGCCAC
CAACTCTGAGTCCAGCACAGTGTCCAGTGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGTCAAGCACAGCCACCAACTCTGAGTCCAGCACACCCCTCAGTGGGCCAGCACAGCACA
GCCACCAACTCTGACTCCAGCACACCCCTCAGTGGGCCAGCACAGCCACCAACTCTGAGTCTAG
CACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCCTCAGTGGGCCA
ACACAGGCCACCAACTCTGGTCCAGTGTGACCTCTGCAGGCTCTGAAACAGCAGCTGACTGG
ATGCACACAACCTCCATAGTCATCTACTGCAGTGAGTGAGGCAAAGCCTGGTGGTCCCTGGT
GCCGTGGAAATCTCCTCATCACCCCTGGTCTGGTTGTGGCGGGCGTGGGCTTTGCTGGC
TCTTCTCTGTGTGAGAAACAGCCTGTCCTGAGAAACACCTTAAACACAGCTGTCTACCACCT
CATGGCCTCAACCATGGCTTGGTCCAGGCCCTGGAGGAAATCATGGAGCCCCCACAGGCCAG
GTGGAGCTTAAGTGGTCTGGAGGAGACCAAGTATCGATGCCATGGAGATGAGCGGGAGGA
ACAGCGGGCCCTGAGCAGCCCCGAACCAAGTGCCGCATTCTCAGGAAGGAAGAGACCTGGCA
CCCAAGACCTGGTTCTCTCATTCACTCCAGGAGACCCCTCCAGCTTGTGAGATCTGAA
AATCTGAAGAAGGTATTCTCACCTITCTGCTTACAGACACTGGAAAGAGAAATACTATAT
TGCTCATTAGCTAAGAAATAACATCTCATCTAACACACAGCACAAAGAGAAAGCTGTGCTTG
CCCCGGGGTGGGTATCTAGCTGAGATGAACTCAGTTAGGAGAAAACCTCCATGCTGGACTC
CATCTGGCATTCAAATCTCCACAGTAAAATCCAAAGACCTCAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKM**Q**KGNVLLM**F**LLLH**L**EAATNSNETSTSANTGSSVISSGASTATNGSSVTSSGVSTATISGS
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESSTPSS
GASTVTVN**G**SSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNDSSTTSSGASTA
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGAGTATNSES
STTSSGASTATNDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTSSGVSTATNSESSTTSSGASTATNDSSTTSSSEASTATNSESSTVSSGISTVTNSES
STTSSGANTATN**G**SSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVVVA
AVGLFAGLFFCVRNLSLRNTFNTAVYHPHGLNHGLGP~~GG~~NHGAPHRPRWSPNWFWRPVSSI
AMEMSGRNNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCCTCGCGTTACGGGATGAATTAAACGGCGGGTCCGCACGGAGGTGTGACCCCTA
CGGAGCCCCAGCTGCCACGCACCCCACACTCGCGTCGCACGGCGTGCCCTGCTTGTACAGGTG
GGAGGCTGAACTATCAGGCTAAAAACAGAGTGGTACTCTCTGGAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGAAAGGGAAATTGTGGTGCTCTGAACCCATGGTCAATT
AACGAGGCAGTTCTAGCTACTGCACGTACTCATAAAGCAGGACTCTAAAGCTTGGAAATCAT
GGTGTATGGAAAGGGATTACTTATACTGACTCTGTTGGGGAGCTTTGGAAAGCATT
TCATGCTGAGTCCCTTTACCTTGATGTTGTAACCCATCTGGTATCGCTGGATCAACAAAC
CGCCTGTGGCAACATGGCTCACCCCTACCTGTGGCATTATTGGAGACCAGTTGGTGTAAAAGT
GATTATAACTGGGATGCATTGTTCTGGAGAAAGAAGTGTCAATTATCATGAACCACCGGACAA
GAATGGACTGGATGTCCTGTGAAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGGTGGCCATGCAGGCTGCTGCCTATAT
CTTCATTCTAGGAAATGGAAGGATGACAAGAGCCATTGCAAGACATGATTGATTACTTTGTG
ATATTCACTGAACCACCTCAACTCCTCATATTCCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCAATTGCTAAAAAAATGGACTTCAGAAATATGAATATGTTTACATCCAAG
AACTACAGGCTTACTTTGTGGTAGACCGTCTAAGAGAAGGTAAAGAACCTTGATGCTGTCCATG
ATATCACTGTGGCGTATCCTCACACATTCCCAATCAGAGAACCTCCCTCAAGGAGACTTT
CCCAGGAAATCCACTTCACGTCCACCGGTATCCAATAGACACCCCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGGCCACAAACGGTGGAGAGAAAGAGAGAGAGGCTGCGTTCTTCTATCAAG
GGGAGAAGAATTTTATTTACCGGACAGAGTGTCAATTCCACCTGCAAGTCTGAACCTAGGGTC
CTTGTGGTCAAATTGCTCTATACTGTATTGGACCTGTTAGCCCTGCAATGTGCCTACTCAT
ATATTGTACAGTCTGTTAAGGGTATTATAATCACCATTGTAATCTTGTGCTGCAAGAGA
GAATATTGGTGGACTGGAGATCATAGAACTTGCATGTTACCGACTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAAATGAGTAAGATTATAAGGTTGCCATGTGAAAACCTAGAGCATATTG
GAAATGTTCTAACCTTCTAACGCTCAGATGCATTGCACTGACTATGTCGAATATTCTTACT
GCCATCATTATTGTTAAAGATATTGCACTTAATTGTGGAAAAATATTGCTACAATT
TTAATCTCTGAATGTAATTGATACTGTGTACATAGCAGGGAGTGTACGGGGTGAATAACTT
GGGCCAGAATATTAAACAATCATCAGGCTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSFGIMVSWKGIYFILTLFWGSFFGSIFMLSP
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM
FLWNCLMRYSYLRLEKICLKLKGVPFGFWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSRSNAFAEKNGLQKYEYVLHPRTGFTFVVDRLREGKNLDAVHDITVA
YPHNIPQSEKHLLQGDFPREIHFHVRYPIDLPTSKEDLQLWCHKRWEKEERLRSFYQGEKNF
YFTGQSVIPPCKSELRLVLLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVFVLQERIFGG
LEIIIELACYRLLKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCACGGCTCCTGCGCCTGAGACAGCTGGCCTGACC
TCCAAATCATCCATCACCCTGCTGTATCTGTTTCATAGTGTGAGATCAACCCACAGGAATA
TCCATGGCTTTGTGCTCATTTGGTCTCAGTTCTACGGAGCTGGTGTCAGGACAGTGGCAAGT
CACTGGACCAGGGCAAGTTGTCCAGGCCTTGGTGGGGAGGACGCCGTGTTCTCTGCTCCCT
TTCCCTGAGACCAGTGCAAGGGCTATGGAAGTGCCTTCTCAGGAATCAGTTCCATGCTGTGGTC
CACCTCTACAGAGATGGGAAGACTGGGAATCTAACAGCAGATGCCACAGTATCGAGGGAAACTGA
GTTTGTGAAGGACTCATTGCAAGGGGGCGTGTCTCTAAGGCTAAAAACATCCTCCTCGG
ACATCGGCTGTATGGGTGCTGGTCAAGTTCCAGATTACGATGAGGAGGCCACTGGAGCTG
CGGGTGGCAGCACTGGCTCACTCCTCTCATTTCCATCGTGGGATATGTTGACGGAGGTATCCA
GTTACTCTGCCTGTCTCAGGCTGGTCCCCAGGCCACAGCCAAGTGGAAAGGTCCAAAGGAC
AGGATTGTCTCAGACTCCAGAGCAAATGCAAGATGGTACAGCCTGTATGATGTGGAGATCTCC
ATTATAGTCCAGGAAATGCTGGAGCATATTGTGTTCCATCCACCTGCTGAGCAGAGTCATGA
GGTGGAACTCAAGGTATTGATAGGGAGAGACGCTTTCCAGCCCTCACCTTGGCCTGGCTTCTA
TTTACTCGGGTTACTCTGTTGCTGTGGTGTGTTGATGGGATGATAATTGTTTCTC
AAATCCAAGGGAAATCCAGGGAAACTGGACTGGAGAAAGCACGGACAGGCAGAATTGAG
AGACGCCGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCGC
TTCTGATCTGAAAATGTAACCCATAGAAAGCTCCCAGGAGGTGCCTACTCTGAGAAGAGA
TTTACAAGGAAGAGTGTGGTGGCTCTCAGGGTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAAATGTAGGGTGGTATGTGGGAGTGTGTCGGGATGACGTAGACAGGGGAAGAAC
ATGTGACTTGTCTCCAACAATGGGTATTGGTCTCAGACTGACAACAGAACATTGTATTC
ACATTCAATCCCCATTATCAGCCTCCCCCAGCACCCCTCTACACGAGTAGGGGTCTTCT
GGACTATGAGGTGGGACCATCTCCTCTCAATACAATGACCAGTCCCTTATTATACCTGC
TGACATGTCAGTTGAAGGCTTGTGAGACCTATATCCAGCATGCGATGTATGACGAGGAAAG
GGGACTCCCATATTCAATGTCCAGTGTCTGGGAT**TGAGACAGAGAACCCCTGCTTAAAGGC**
CCACACACAGACCCAGACACAGCCAAGGGAGAGTGTCCCGACAGGTGGCCCCAGCTTCT
CCGGAGCCTGCGCACAGAGACTCACGCCCTACTCTCCTTAGGGAGCTGAGGTTCTCTGCC
TGAGCCCTGCAGCAGCGGCACTCACAGCTCCAGATGAGGGGGATGGCCTGACCTGTGGAG
TCAGAACCCATGGCTGCCCTGAAGTGGGAGCGGAATAGACTCACATTAGTTAGTTGTAAAA
CTCCATCCAGCTAACGATCTGAACAAGTCACAACCTCCAGGCTCTCATTTGCTAGTCACGG
ACAGTGATTCTGCTCACAGGTGAAGATTAAGAGACAAACGAATGTGAATCATGCTTGCAAGTT
TGAGGGACAGTGTGCTAATGATGTGTTTATATTACATTTCACCATAAACTCTGTT
TGCTTATTCCACATTAATTACTTTCTCTATACCAAAATCACCCATGGAATAGTTATTGAACACC
TGCTTGTGAGGCTAAAGAATAAGAGGAGGTAGGATTTCACTGATTCTATAAGCCCAGCAT
TACCTGATACCAAAACAGGAAAGAAAACAGAAGAAGAGGAAGGAAACTACAGGTCCATATCC
CTCATTAAACACAGACACAAAAATTCTAAATAAAATTAACTAAACAAATTATTTA
AAGATGATATATAACTACTCAGTGTGGTTGTCCCACAAATGCAGAGTTGGTTAATATTAAAT
ATCAACCAGTGAATTCAAGCACATTAATAAGTAAAAAGAAAACCATAAAAAAAAAAAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH
LYRDGEDWESKQMPQYRGRTEFKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDDEATWELR
VAALGSLPLISIVGYVDGGIQLLCLSSGWFPOPTAKWKGPGQDLSSDSRANADGYSLYDVEISI
IVQENAGGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALCGVVMGMIIVFFK
SKGKIQAELDWRRKHGQAEELRDARKHAVEVTLDPETAHPKLCVSDLKTTHRKAQEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNNGWYVGVCRDDVDRGKNNVTLSPNNGYWVLRLTTEHLYFT
FNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLRPYIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTCATTGCTGGTGGCAATGATGTATCGGCCAGATGTTGAGGGCTAGGAAAAGAG
TTTGGTGGAACCTGGTTATCGCCTCGTCATCTCATATCCCTGATTGTCCTGGCAGTGTGCATTGGA
CTCACTGTTCAATTATGTGAGATATAATCAAAGAACCTACAATTACTATAGCACATTGTCATTACAC
TGACAAACTATATGCTGAGTTGGCAGAGAGGGCTCTAACAAATTACAGAAATGAGCCAGAGACTTGAAT
CAATGGTAAAAATGCATTTATAAATCTCCATTAAGGAAGAATTGTCAAGTCTCAGGTTATCAAGTTC
AGTCAACAGAACATGGAGTGTGGCTCATATGCTGTTGATTGTAGATTCTACTCTAGGATCCTGA
AACTGTAGATAAAATTGTCACATTGTTTACATGAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAG
ATCCTCACTCAGTTAAAATTAAAAAAATCAACAAGACAGAACAGACAGCTATCTAAACCATTGCTGGGA
ACACGAAGAAGTAAAACCTAGGTCAAGTCAGGATCGTGGTGGACAGAAGTAGAAGAGGGTGAATG
GCCCTGGCAGGCTAGCCTGCAGTGGGATGGAGTCATCGCTGAGCAACCTTAATTAGCCACATGGC
TTGTGAGTGCTGCTCACTGTTTACAACATATAAGAACCTGCCAGATGGACTGCTTCCTTGAGTAACA
ATAAAACCTCGAAAATGAAACGGGTCTCGGAGAATAATTGTCATGAAAATACAAACACCCATCACA
TGACTATGATATTCTCTTGCAAGCAGCTTCTAGCCCTGTTACACAAATGCAAGTACATAGAGTTGTC
TCCCTGATGCATCCTATGAGTTCAACCAAGGTGATGTGATGTTGTGACAGGATTGGAGCACTGAAAAT
GATGGTTACAGTCAAATCATCTCGACAAGCACAGGTGACTCTCATAGACGCTACAACCTGCAATGAACC
TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTAGAAGGAAAACAGATGCAT
GCCAGGGTGAECTGGAGGACACTGGTAGTTCAAGATGCTAGAGATATCTGGTACCTGCTGGAATAGTG
AGCTGGGAGATGAATGTGCAAAACCAACAAGCCTGGTGTATACTAGAGTTACGGCCTGCGGGACTG
GATTACTCAAAAATGGTATCTTAAGAGACAAAGCCTCATGGAACAGATAACTTTTTTGTTTG
GGTGTGGAGGCCATTAGAGATAACAGAATTGGAGAAGACTGCAAAACAGCTAGATTGACTGATCTCA
ATAAAACTGTTGCTTGATGCATGTTCTAGAAGTTGACTGTTGACATAATTGTAAT
GATCAACTCTGTCACTGTGAGCAATAGTTGAAACTTATGTACATAGAGAAATAGATAACAAATTAC
ATTACAGCCTGTATTCAATTGTTCTAGAAGTTGACTGTTGACATAATTGTAAT
GCATATATAACATTGAAGCACTCCTTCTCAGTTCCAGCTGTTCCGCACGTAAGCATCCTGCTCTGCCA
TCAAGGTGCAGAACAGGAGTGAAGAAAATAAGAAGAAAAATCCCCTACATTGACAGAA
AAGTATTAGGTGTTCTAGTGAATAATTAGAAATGATCATATTGAAAGGTCAAGCAAAGACA
GCAGAACATCAACTCATCATTAGGAAGTATGGAACTAAGTTAAGGAAGTCCAGAAAGAAGCCAAG
ATATATCCTTATTCATTCCAAACAACACTATGATAATGTGAAGAAGATTCTGTTTGACCT
ATAATAATTACAAACTTCATGCAATGACTTGTCAAGCAAATTAAAGCAAATTATTAAACATTG
TTACTGAGGATGTCAACATATAACAAATAAAATATAACCTTACCCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTGHYVRYNQKKTYNYYSTLSFTTDKLY
AEFGREASNNFTEMSQRLESMVKNAYKSPLREFVKSQVIKFSQQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATLVSAAHCFTTYKNPARWTASFGVTIKPSKMKGRL
RRIIVHEKYKHPSHDYDISLAELSSPVPTNAVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
QNHLRQAQVTLIDATTNEPQAYNDAITPRMLCAGSLEGKTDACQGDGGPLVSSDARDIWYLAG
IVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCGCGAAGAAGTTCCCTG
CCCCGATGAGCCCCCGCCGTGCGTCCCCGACTATCCCCAGGCGGGCGTGGGGCACCGGGCCCAGC
GCCGACGATCGCTGCCGTTTGCCCTGGGAGTAGGATGTGGTGAAGGTCTGCGCTGCCCTGCTCACGCCCTCAA
ACGGGGCTCACAATGCCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCCTGCTCACGCCCTCAA
TCTGCTCTTTGGTTAATGTCCATCAGTGTGGCAGTTCTGCTGGATGAGGGACTACCTAA
ATAATGTTCTCACTTAACGTGAGAACGAGGGTAGAGGAAGCAGTCATTGACTTACTTCT
GTGGTTCATCCGGTCAATTGCTGTTGCTGTTCTATCATTGTTGGGATGTTAGGATATTG
TGGAACGGTAAAAGAAATCTGTGCTCTTGCACTGGTACTTTGAAGTTGCTTGTCAATTCT
GTGTAGAACTGGCTTGTGGCGTTGGACATATGAACAGGAACCTATGGTCCAGTACAATGGTCA
GATATGGTCACTTGAAGGCCAGGATGACAATTATGGATTACCTAGAATATCGTGGCTTACTCA
TGCTTGAATTTCAGAGAGGTTAAGTGCTGTTGGAGTAGTATATTCACTGACTGGTTGG
AAATGACAGAGATGGACTGGCCCCCAGATTCTGCTGTGTTAGAGAATTCCCAGGATGTTCAAA
CAGGCCACCAGGAAGATCTCAGTGACCTTATCAAGAGGGTGTGGGAAGAAAATGTAATTCTT
TTTGAGAGGAACCAACAACTGCAGGTGCTGAGGTTCTGGGAATCTCAATTGGGTGACACAAA
TCCCTGGCATGATTCTCACCATTACTCTGCTGGGCTGTATTATGATAGAAGGGAGCCTGG
ACAGACCAAATGATGTCCTGAAGAATGACAACCTCTCAGCACCTGTCATGTCCTCAGTAGAACT
GTTGAAACCAAGCCTGCAAGAATCTTGAAACACACATCCATGGCAAACAGCTTAATACACACT
TTGAGATGGAGGTTAAAAAAAAATGTCACAGAAGAAAACCACAAACTGTTTATTGGACT
TGTGAATTGAGTACATACTATGTTGCTGAAATATGTTAGAAATAAAATGTTGCCATAAAA
TAACACCTAACATACTATTCTATGCTTAAATGAGGATGAAAAGTTCATGTCATAAGTC
ACCACCTGGACAATAATTGATGCCCTAAATGCTGAAGACAGATGTCATACCACTGTTAGCC
TGTGTATGACTTTACTGAACACAGTTATGTTGAGGCACTGGTTGATTAGCATTCCGCA
TCCATGCAAACGAGTCACATATGGGGACTGGGCCATAGTAAAGGTTGATTTACTTCTACCAA
CTAGTATAAAAGTAACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATT
CTCAGCGATCTATTCTCTGATGCTAAATAATTATCAGAAAACCTTCAATATTGGTACT
ACCTAAATGTTGATTTGCTGGTTACTAAATATTCTTACCACTAAAGAGCAAGCTAACACAT
TGTCTTAAGCTGATCAGGGATTGGTATATAAGTGTGTTAAATCTGTATAATTCACTGAT
TTCAGTTCTGATAATGTTAAGAATAACCATTGAAAAGGAAATTGTCCTGTATAGCATCATT
ATTTTACGCTTCTGTTAATAAGCTTACTATTCTGCTCTGGCTTATATTACACATATAAC
TGTATTAAATCTAACACTAATTGAAAATTACCACTGTTAGGAAATCTGATCATT
AGAATGTTAGTCTGGCTTTAGGAAGTATTAATAAGAAAATTGCACTAACTTAGTTGATTGAGA
AAGGACTTGTATGCTGTTTCTCCAAATGAAGACTCTTTGACACTAAACACTTTAAAAAA
GCTTATTTGCCTCTCCAAACAAGAAGCAATAGTCTCAAGTCATATAAAATTCTACAGAAA
TAGTGTCTTTCTCCAGAAAATGCTGAGAATCATTAAACATGTGACAATTAGAGATT
CTTTGTTTATTCTACTGATTAATATACTGTCGCAAATTACACAGATTATTAATTTTACAA
GAGTATGTTATTGAAATGGAAAAGTGCATTACTGTTAGTATTTGTTATTTGTTAT
TTCTCAGAATATGAAAGAAAATTAAATGTCATAAAATTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWIIMSISVLAVSAWMRDYLNNVLTLTAE
TRVEEAVILTYFPVVHP
VMIAVCCFLIIVGMLGYCGTVKRNLILLAWYFGSLLVI
FCVELACGVWTYEQELMVPVQWSDMVT
LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTEMDWPPDSCCVREFPGCSKQAHQ
EDLSDLYQEGCGKKMYSFLRGTKQLQVIRFLGISIGVTQILAMILTITLLWALYYDRREPGTDQM
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGGACACCTTATCCCACTCATCCTCATCCTCTTCCTGTATAAAGCCCTACCAGTGCT
GATAAAGCTTTCTCGTAGAGGCCTAGAGGCCCTAAAAAAAGTGCCTGAAAGAGAAGGGGACAAGGAACA
CCAGTATTAAGAGGATTTCAGTGTTCCTGGCAGTTGGTCCAGAAGGATGCCCCATTCTGCTTCACCTG
CCTCTCATCACAGGCACCTCCGTGTCACCCGTGCCCTAGATCCTGTCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACCAGTGGATGAGTCAGTCAGGTCTCTCTATGTGACAACCATGTGAATGGGAG
TGGTACCACTTCAGGGCATGGCGGAGATGCCATGCCTACCTCTGCATACAGAAAACACTGTGGAACCCA
CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAGGCGACGGCATTGTGCAACGCCAGGCTGTGCCAGCT
TCAATGGGAACTGCTGTCCTGGAACACCACGGTGGAAAGTCAGGCTTGCCCTGGAGGCTACTATGTGATCGT
CTGACCAAGCCCAGCGTCTGCTTCACGTCTACTGTGGTCAAGGACTCTGCGACAGGGACTGCCATGG
CAGCTGCTCAGATAACCAGCGAGTGACATGCCCTCAGGAACCTGCTAGGCCCTGACAGGCAGACATGCTTG
ATGAAAATGAATGTGAGCAAACACGGTGGCTGCACTGAGATCTGTGAAACCTCAAAACTCTACCGCTGT
GAGTGTGGGTTGGCGTGTGCTAAGAAGTGTGAGACTGTGAGACGTTGAGATGCCACAATAACAA
TGGTGGCTGAGCCACTCTTCCTGGATCTGAGAAAAGGCTACCGTGTGAAATGTCCCCGGGCTGGTGT
CTGAGGATAACCACACTTGCCTGGAGACTCTTCCTGACCAACACCTCCGAGGAGTGTCAACGGCACCCATGTCAACAT
CCTCTCTCTCAAGACATGTGGTACAGTGGTGTGGTGAATGACAAGATTGTCAGGCTGCAACCTCGTGA
CAGGTCTACCCAAGCAGACCCGGGGAGCAGCGGGACTTCATCATCCGAACAGCAGCTGCTGATCCGGTG
ACCTGCGAGTTCCACGCCGTACACCATTCTGAAGGATACGTTCCAACTTCGAAACTCCCCACTGAAAT
CATGACCGAAATCATGGATCTTCCATTCACTCTGGAGATCTCAAGGACAATGAGTTGAAGAGCCTTAC
GGGAAGCTCTGCCACCCCTCAAGCTCGTACTCCCTACTTGGCATTGAGGCCGTGGTGCACGTGAGCGGC
TTGGAAAGCTTGGTGGAGAGCTGCTTGCACCCACCTCCAAGATCGACGAGGCTCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTCAAGTGTGACTCGTAAAGCAGTACACATCCGGGATCACCTAGCAAAGCACTTCCAGG
TCCCTGTTCAAGTTGTGGCAAAGACCACAAGGAAGTGTTCGACTGCCGGTTCTGTGTGGAGTG
TTGGACGAGCGTCCCGCTGTGCCACCCGAGGCTGCCACCGCGAATCGTGTGGCAGGAGGAGGACTCAGC
CGGTCTACAGGGCCAGACGCTAACAGCGGCCGATCCGACTGGGAGGACTTAGTCGTAGCCATACCTC
GAGTCCCTGCATTGGACGGCTCTGCTCTTGGAGCTCTCCCCCAGGCCCTTAAGAACATCTGCCAACAGC
TGGGTTCAAGACTCACACTGTGAGTCAGACTCCACCAACTCACTCTGATTCTGGTCCATTCAAGTGGCA
CAGGTCAAGCAGCTGCTGAACAATGTGGCTGGGTGGGTTTCATCTTCTAGGGTTGAAAACAAACTGTCCA
CCCAGAAAGACACTACCCCTATTCCCTCATTTCTTCTACACTTAAACACCTCGTGTATGGTGCATCAGAC
CACAAAATCAGAAGCTGGTATAATATTCAAGTTACAAACCTAGAAAATTAACAGTTACTGAAATTATGA
CTTAAATACCAATGACTCCTAAATATGTAATTAGTTACCTGAAATTCAATTCAATGCAGACTAA
TTATAGGAAATTGGAAGTGTATCAATAAACAGTATATAATT

FIGURE 110

MPPFLLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPENHCGTHAPWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPGGYYVRLTKPSVCFHV
YCGHFYDICDEDCHGCSDTSECTCAPGTVLGPDRQCFDENECQNNGGCSEICVNLKNSYRCECGVGRV
LRSRGKTCEDVEGCHNNNGGCSHSCLGSEKGYQCECPRLVLSEDNHTCQVPVLCKSNAIEVNIPRELVGG
LELFILTNTSCRGVSNGTHVNILFSLKTCGTVVDDVVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFITLEIFKDNEFEEPYREALPTLKLRSLSLYFGIEPVHV
SGLESILVESCATPTSKIDEVLKYYLIRDGCVSSDSVKQYTSRDHLAKHFQVPVFVKGKDHEVFLHCRV
LVCGVLDERSRCAQGCHRRMRRGAGGEDSAGLQQQTLTGGPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTGCTCAGCGACAAGGATGCTGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG
GCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGCCAGCCAGCACGTGTTGGGAGGCC
CCTGCTGCCTTGGGTGACAATCTAGCTCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT
TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCGTATCCCC
ATGGAGACCTTCAGAAAGGTGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGT
GGTTGTCTCATCAAGGTGATTCTGGATAAATACTACTTCCTCTGCCAGCCTCTCCACTTCATCCGA
GGAAGCAGCTGTGTGACGGAGAGCTGGACTGCCCTGGGGAGGACGAGGAGCACTGTGTCAGAGCTTC
CCCGAAGGGCCTGCAGTGGCAGTCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
AGGGAACTGGTTCTCTGCCTGTTGACAACCTCACAGAAGCTCGCTGAGACAGCCTGTAGGCAGATGG
GCTACAGCAGAGCTGTGGAGATTGCCAGACCAAGGATCTGGATGTTGAAATCACAGAAAACAGCCAG
GAGCTTCGCATGCGGAACCTCAAGTGGCCCTGTCCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCC
TGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGTGGGGAGGAGGCCCTGTGGATTCTTGGCTTGGCAGG
TCAGCATCCAGTACGACAAACAGCACGCTGTGGAGGGAGCATCCCTGGACCCCCACTGGGCCTCACGGCA
GCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCCAGGCTCAGACAAACTGGCAG
CTTCCCACCCCTGGCTGTGGCAAGATCATCATTAATTCAACCCATGTACCCAAAGACAATGACA
TCGCCCTCATGAAGCTGCAAGTCCCACACTTCTCAGGCACAGTCAGGCCATCTGTGCCCTTCTT
GATGAGGAGCTCACTCCAGCCACCCACTCTGGATCATTGGATGGGCTTACGAAGCAGAACGGGAA
GATGTCAGACATACTGTCAGGGCTCAGTCAGGTCATTGACAGCACACGGTCAATGAGACGATGCGT
ACCAGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGAAGGGGGTGTGGACACCTGCCAGGG
GACAGTGGTGGGCCCTGATGTACCAACTGACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT
CTGGGGGGCCCGAGCACCCAGGGAGTACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT
GGAAGGCTGAGCTGTAATGCTGCCCTTGCAGTGCTGGAGGCCCTGCCCTGCCACCT
GGGGATCCCCAAAGTCAGACACAGAGCAAGAGTCCCTGGTACACCCCTCTGCCACAGCCTCAGCAT
TTCTGGAGCAGCAAAGGCCCTCAATTCTGTAAAGAGACCCCTCGCAGCCAGAGGCCAGAGGAAGTCA
GCAGCCCTAGCTGCCACACTTGGTCTCCAGCATCCAGGGAGAGACACAGCCCAGTGAACAAGGTCT
CAGGGGTATTGCTAACGCAAGAAGGAACCTTCCACACTACTGAATGGAAGCAGGCTGTCTGTAAAAGCC
CAGATCACTGTGGCTGGAGAGGAGAAGGAAAGGGTCTGCCAGCCCTGCGTCTTCAACCCATCCCCAA
GCCTACTAGAGCAAGAACCAAGTTGTAATATAAAATGCACTGCCCTACTGTGGTATGACTACCGTTACCT
ACTGTTGTCATTGTTATTACAGCTATGCCACTATTATAAGAGCTGTGTAACATCTCTGGCAAAAAAAA
AAAA

FIGURE 112

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIP III ALLSLASIIIVVVLIKVILDKYFLCG
QPLHFIPRKQLCDGELDCPLGEDEEHCVKS FPEGPAVAVRLSKDRSTLQVLD SATGNWFSACFDN
FTEALAE TACRQM GYSRAVEIGPDQDLDVVEITENSQELMRNSSGPCLSGSLVSLHCLACGKSL
KTPRVVGEEASVDSWPWQVS IQYDKQHVC CGSILDPHWVL TAAHCFRKHTDVFNWKVRAGSDKL
GSFPSLAVAK IIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICL PFFDEELTPATPLWIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDT CQGD SGGPLMYQS
DQWHVVGIVSWG YGC GG PSTPGVYTKV SAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCCTGGTCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTCTGCCCTCTGAACAACCTTTAGCAACTAAAAAGCCACAGGAGT
TGAACTAGCTAGGATTCTGACTATGCAGCTGGCTAGTGCTCCTACTCCTACCTACATTAAAATC
TGTCCCCCTGTTCTGTAACTAGCCTTACCTCCTAACACAGAGGATCTGTCAGTGGCTCT
GGCCCAAACCTGACCTCACTCTGGAACGAGAACAGAGGTTCTACCCACACCGTCCCCTGAAG
CCGGGGACAGCCTCACCTGCTGCCCTCTCGCTGGAGCAGTGCCTCACCAACTGTCTCACGTCT
GGAGGCACTGACTCGGGCAGTGCAGGTAGCTGAGCCTCTGGTAGCTGGCTTCAAGGTGGC
CTTGCCCTGGCGTAGAAGGGATTTGACAAGCCGAAGATTCTAGGCATGGCTCCACTGCC
AGGCATCAGCCTTGCTGTAGTCATCACTGCCCTGGGCCAGGACGGCCGTGGACACCTGCTCA
GAAGCAGTGGGTGAGACATCACGCTGCCGCCATCTAACCTTTCATGTCCTGCACATCACCTG
ATCCATGGCTAATCTGAACACTGTCCCAAGGAACCCAGAGCTTGAGTGGCTCAGACC
CAGAAGGGTCTGCTTAGACCACCTGGTTATGTGACAGGACTTGCAATTCTGGAACATGAGG
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAAGGAACCTTGTCACAAATTATGGGTAGAAAAGATG
GAGGTGTTGGTTATCACAAGGCATCGAGTCTCCTGCATTCACTGGACATGTGGGGAAAGGGCTG
CCGATGGCCATGACACACTCGGACTCACCTCTGGGCCATCAGACAGCCGTTCCGCCCGAT
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCATGCTCTCATCAGCCAGGCAGGCCAAA
TCTGCGATCACCAGCCAGGGCAGCCGCTGGAGGAGCAAGCAAAGTGACCAATTCTCCCTCC
CTCCTCCCTGAGAGGCCCTCTATGTCCTACTAAAGCCACAGCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTGGCCAGGAGGTAGCAAGGCCAGGAGCTGATCAGAAGGGCTGCT
GTGCGAACACGGAAATGCCCTCAGTAAGCACAGGCTGCAAAATCCCAGGCAAAGGACTGTGTGG
CTCAATTAAATCATGTTCTAGTAATTGGAGCTGTCCCAAGACCAAGGAGCTAGAGCTTGGTT
CAAATGATCTCCAAGGCCCTTATACCCAGGAGACTTGATTTGAAATTGAAACCCCAAATCCA
AACCTAAGAACCGAGGTGCAATTAGAAATCAGTTATTGCCGGTGTGGCTGTAATGCCAACAT
TTTGGGAGGCCAGGGCTAGATCACCTGAGGTAGGAGCTCAAGACAGCCTGGCCAACATGG
TGAAACCCCTGTCTACTAAAAATACAAAAAAACTAGCCAGGCATGGGTGTGCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAAACCTGGAGGTGAAGGAGGCTGAGACA
GGAGAACACTCAGCCTGAGCAACACAGCGAGACTCTGTCAGAAAAATAAAAAAGAATTA
TGGTTATTGTAA

FIGURE 114

MLWWLVLLLLPTLKVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRD

Signal peptide:

amino acids 1-15

FIGURE 115

CAGCAGTGGTCTCTAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTACATTCTAAATGCAGAAGCTTTAAATCCAAGAAA
ATATGTAATCACTTAAGATTGTGGACTGGTGGTATCCTGGCCCTAACCTAATTGTCC
GTTTGCCCCAGCAAGCACTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACT
TCTACAGCAATGGAGAGAAGAAGAAGATTACATGAAATTGATCCTGTGACCAGAACTGAAATA
TTCAGAAGCGGAAATGGCACTGATGAAACATTGAAAGTGCACGACTTAAAAACGGATAACTGG
CATCTACTTCGTGGTCTTCAAAAATGTTTATCAAAACTCAGATTAAAGTGATTCTGAATT
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCAACTTTCTTGAAACAGTCAGTG
ATTTGGTCCCAGCAGAAAAGCTATTGAAAACCGAGATTCTTAAAAATTCCAAAATTCTGGA
GATTTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAAATATCAGTTCTGAGTTACAAG
ACTTTGAGGAGGGAGGAGAAGATCTTCACTTCCTGCCAACGAAAAAAAAGGGATTGAACAAAAT
GAACAGTGGGTGGTCCCTCAAGTGAAGTAGAGAAGACCCGTACGCCAGACAAGCAAGTGAGGA
AGAACTCCAATAATGACTATACTGAAAATGGAATAGAATTGATCCCATGCTGGATGAGAGAG
GTTATTGTTGTATTACTGCCGTGAGGCAACCGCTATTGCCGCCGCTGTGAACCTTACTA
GGCTACTACCCATATCCATACTGCTACCAAGGGAGGACGAGTCATCTGCTGTCACTCATGCC
TAACCTGGTGGGTGGCCCGATGCTGGGGAGGGTCAATAGGAGGTTGAGCTCAAATGCTAAAC
TGCTGGCAACATATAATAATGATGCTATTCAATGAATTCTGCCTATGAGGCATCTGGCC
GGTAGGCCAGCTCTCCAGAATTACTGTAGGTAAATTCTCTTCAATGTTCTAATAAAACTCTACA
TTATCACCAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFSKICKSLKICGLVFGILALTLLIVLFWGSKHFWPEVPKKAYDME
HTFYSNGEKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTGIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEIITTFEQSVIWVPAEKPIENRDFLKNSKILEICDNVTMYWINPTLISVSE
LQDFEEEEDLHFPAANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMED
ERGYCCIYCRRGNRYCRRVCEPLLGYYPPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-
242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCTCAGGAGCGCGTTAGCTTCACACCTCGGCAGCAGGAGGGCGCAGCTCTCGCAGGCGCA
GGCGGGCGGCCAGGATCATGTCCACCACATGCCAAGTGGTGGCCTCCTGTCCATCCTGGGCT
GGCGGCTGCATCGCGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCCGTACCT
CCGTGTTCACTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCC
TATTCACCATCCTGGACTTCCAGCCATGCTGCAGGCAGTGCAGCCCTGATGATCGTAGGCATCGTCC
GGGTGCCATTGGCCTCCTGGTATCCATCTTGCCCTGAAATGCATCCGATGGCAGCATGGAGGACTCTG
CCAAAGCCAACATGACACTGACCTCCGGGATCATGTTCAATTGTCATTGTCAGGTCTTGTGCAATTGCTGGAGTG
TCTGTGTTGCCAACATGCTGGTACACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGG
GATGGTGCAGACTGTTCAGACCAGGTACACATTGGTGCAGCTGTGAGGCTGGTCGCTGGAGGCC
TCACACTAATTGGGGGTGTGATGATGTGCATGCCCTGCCGGGCCTGGCACAGAAGAAACCAACTACAAA
GCCGTTCTTATCATGCCCTCAGGCCACAGTGGCCTACAAGCCTGGAGGCTCAAGGCCAGCACTGGCTT
TGGGTCCAACACCAAAACAAGAAGATAACGATGGAGGTGCCGCACAGAGGACGAGGTACAATCTTATC
CTTCCAAGCACGACTATGTGTAATGCTCTAACAGCTCTCAGCACGGCGGAAGAAACTCCGGAGAGCTCA
CCCCAAAAACAAGGAGATCCCCTAGATTTCTGCTTTGACTCACAGCTGGAAGTTAGAAAAGCCT
CGATTTCATCTTGAGAGGCCAAATGGCTTAGCCTCAGTCTGTCTCAAATATTCCACCATAAAACA
GCTGAGTTATTATGAATTAGAGGCTATAGCTCACATTCAATCCTCTATTCTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGTTTAATCTCTCTCACATTGATGATTTAGACAGACTCCCCCTC
TTCCCTCTAGTCATAAAACCCATTGATGATCTATTCCCAGCTTATCCCAAGAAAATTGGCTAGTAATAA
GAGTAGACCCAAAGATGTTATTCTGCTGTTGAATTGCTCCCCACCCCAACTGGCTAGTAATAA
ACACTTACTGAAGAAGCAATAAGAGAAAGATATTGTAATCTCTCCAGCCCATGATCTGGTTCTT
ACACTGTGATCTAAAGTTACCAACCAAAAGTCATTTCAGTTGAGGCAACCAACCTTTCTACTGCTG
TTGACATCTTCTATTACAGCAACACCATTCTAGGAGTTCTGAGCTCTCCACTGGAGTCCTTTCTGT
CGCGGGTCAGAAATTGCTCTAGATGAATGAGAAAATTATTTTTAATTAAAGTCCTAAATATAGTAA
ATAAAATAATGTTAGAAAATGATAACTATCTCTGAAATAGCCTCACCCCTACATGTGGATAGAAG
GAAATGAAAAAAATAATTGCTTGACATTGCTATATGGTACTTGTAAAGTCATGCTTAAGTACAAATTCC
ATGAAAAGCTCACACCTGTAATCCTAGCATTGGAGGCTGAGGAGGAAGGATCACTTGAGGCCAGAAGT
TCGAGACTAGCCTGGCAACATGGAGAAGCCCTGTCTCACAAACACAGAGGAGAAAATCAGCCAGTC
TGGTGGCATAACACCTGTAAGTCCCAGCATTCCGGAGGCTGAGGTGGAGGATCACTTGAGGCCAGGGAGGT
TGGGGCTGCAGTGAGCCATGATCACACCAACTGCACCCAGGTGACATAGCGAGATCCTGTCTAAAAA
ATAAAAAAAATAATGGAACACAGCAAGTCCTAGGAAGTAGGTTAAAACTAATTCTTAA

FIGURE 118

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLWRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYHASGHHSVAYKPGGFKA^STGFGSNTKNKKIYDGGARTEDEVQSYP^SSKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAAACTGTTCTCTCTGTGGCACAGAGAACCTGCCTCAAAGCAGAAGTAGCAGTCGGAGTCC
AGCTGGCTAAAATCATCCCAGAGATAATGGCAACCCATGCCTAGAAATCGCTGGCTGTTCTG
GTGGTGTGGATGGTGGCACAGTGGCTGTCACTGTCACTGCCTCAGTGGAGAGTGTGGCCTTCATT
GAAAACAACATCGTGGTTTGAAAACCTCTGGGAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
CATCAGGATGCAGTGC~~AAA~~ATCTATGATTCCCTGCTGGCTTTCTCCGGACCTACAGGCAGCCAGAG
GACTGATGTGTGCTGCTCCGTATGTCCTCTGGCTTATGATGCCATCCTGGCATGAAATGC
ACCAGGTGCACGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGATCATCTTCAT
CATCACGGCATGGTGGTGCATCCCTGTGAGCTGGGTGCCATGCCATCATCAGAGATTCTATA
ACTCAATAGTGAATGTTGCCAAAACGTGAGCTGGAGAAGCTCTACTTAGGATGGACCACGGCA
CTGGTGCATGGTGGAGGAGCTGTCCTGCTGCCCTTGTGCAACGAAAAGAGCAGTAGCTA
CAGATACTCGATACCTCCCACCGACAACCCAAAAAGTTATCACACCGGAAAGAAGTCACCGAGCG
TCTACTCCAGAAGTCAGTATGTGTAGTTGTATGTTTAACTTACTATAAGCCATGCAAATG
ACAAAAATCTATATTACTTCTCAAATGGACCCAAAGAAACTTGTGATTACTGTTCTTAACGCCT
AATCTTAATTACAGGAACGTGCATCAGCTATTGATTACTATAAGCTATTCAGCAGAAATGAGATA
TTAAACCAATGCTTGATTGTTCTAGAAAGTATAGTAATTGTTCTAAGGTGGTCAAGCATACTA
CTCTTTTATCATTACTTCTCAAATGACATTGCTAAAGACTGCATTATTACTACTGTAATTCTCC
ACGACATAGCATTATGTACATAGATGAGTGTAAACATTATCTCACATAGAGACATGCTATATGGT
TTTATTTAAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTCAAGGAA
ATCATGGATAGGGTGAAGAAGGTTACTATTGTTAAAACAGCTTAGGGATTATGCTCTCCA
TTTATAATGAAGATTAAGGCTTAAATCAGCATTGTAACAGGAAATTGAATGGCTTCTGATAT
GCTGTTTTAGCCTAGGAGTTAGAAATCCTAACCTCTTATCCTCTCCAGAGGCTTTTT
CTTGTGTATTAAATTAAACATTAAACATTAAACGAGATATTGTCAGGGCTTGCATTCAAACGCTT
TTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAGTGTGTTTAGGAAAGT
AAAATATTTGTTGTATTGAAGAAGAATGATGTCATTGACAAGAAATCATATGTATGGAT
ATATTTAATAAGTATTGAGTACAGACTTGTGAGGTTCTCAATATAAAATAAAAGAGCAGAAAATA
TGTCTGGTTTCATTGCTTACCAAAAAACACAAACAAAAAGTTGTCCTTGAGAAACTCACCT
GCTCCTATGTGGGTACCTGAGTCAAATTGTCATTGTTCTGTGAAAATAAATTCCCTTCTGTA
CCATTCTGTTAGTTACTAAATCTGAAATACTGTGTTCTGTTATTCAAATTGATGAA
ACTGACAATCCAATTGAAAGTTGTCGACGTCTGTCTAGCTAAATGAATGTGTTCTATTGCTT
TATACATTATATAAATTGACAITTTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTAVTVMPQWRVSIFIENNIVVFENFWEGLWMNCVRQANIRMOCK
IYDSLLALSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG
MVVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCNEKSSSY
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGCGGGTGAAGGCGCATTGATGCAGCCGTGGCGGCCCTGGAGCGCGGGAG
CCAGACGCTGACCACGTTCTCTCCCTCGTCTCCCTCCAGCTCCGCGCTGCCCGCAGCC
GGGAGCCATGCGACCCCAGGGCCCCGCGCCCTCCCGCAGCGGCTCCGCGCTCCGCTCC
TGCTGCTGCAGCTGCCCGCGCCGTGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAAAGGCCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGCCAATGTTATTCCGGTACACCTGGATCCCAGGTGGATG
GATTCAAAGGAGAAAAGGGGAATGTCAGGGAAAGCTTGAGGAGTCTGGACACCCAACTAC
AAGCAGTGTTCATGGAGTTCAATTGAATTATGGCATAGATCTGGAAAATTGCGGAGTGTACATT
TACAAAGATGCGTTCAAATAGTGCTCTAACAGAGTTGTTCACTGGCTACTTCGGCTAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTCACATTCAATGGAGCTGAATGTTCAAGGACCTTCCC
ATTGAAGCTATAATTATTGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAAATTATCG
CACTTCTCTGTGGAAGGACTTGTGAAGGAATTGGTGTGGATTAGTGGATGTTGCTATCTGG
TTGGCAGTTGTCAGATTACCAAAAGGAGATGCTTCACTGGATGGAATTCACTGGCTTCG
ATTATTGAAGAACTACCAAAAAATGCTTAATTTCATTTGCTACCTCTTTTATTATGCC
TTGGAATGGTTCACTTAAATGACATTAAATAAGTTATGTATACATCTGAATGAAAAGCAAAG
CTAAATATGTTACAGACCAAGTGTGATTCACACTGTTAAATCTGATTATTCACTGG
CTTCAACCTATAATTGGAATTGTTGGTCTTTGTTCTCTTAGTATAGCATTGTT
AAAAAATATAAAAGCTACCAATTGTTACAATTGTAAATGTTAAGAATTGTTATATCTGT
TAAATAAAATTATTCCAACA

FIGURE 122

MRPQGPAA SPQR LRG LLLL LQL PAPSSASE I PKG KQKA QLR QREVVD LYNG MCL QGPAG VPGR
DGSPGANV IPGTPGIPGRDGFKGEKGECLRESFEESWTPNYKQCSWSSLNYGIDL GKI AECTFTK
MRSNSALRVL FSGSLRLKCRNACCQRWYFTFNGAEC SGPLPIEAI IYLDQGSPEMNSTINI HRTS
SVEGLCEGIGAGLVDVAIWVGTCSDYPKG DASTGWNSVSRIII EELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGTACGGGCTCCTGCCTCTGGCTCCAACG CAGCTGTGGCTGAA
CTGGGTGCTCATCACGGGAACTGCTGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
ATTGCCCTGGAAGAATACATCATGTTTICGATAAGAAGAAATTGTAAGGATCCAGTTTTTTTA
ACCGCCCCCTCCCCACCCCCCAAAAAACTGTAAGATGCAAAACGTAATATCCATGAAGATCC
TATTACCTAGGAAGATTTGATTTGCTGCGAATGGGTGTTGGATTATTGTTCTGGAG
TGTTCTCGTGGCTGGCAAAGAATAATGTCAAAATCGGTCCATCTCCAAGGGTCCAATT
TCTTCCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGCTGTCATGCAACTG
GCCCTAAGCCAAAGCAAAGACCTAAGGACGACCTTGAACAATACAAGGATGGGTTCAATG
TAATTAGGCTACTGAGCGGATCAGCTTAGCACTGGTTAGCCCCACTGTCTACTGACAATG
CTTTCTCTGCCGAACGAGGATGCCCTAAGGGCTGAGGTGTAAGGCAAATGGTATTGTA
ATCTCAGAAATTACAGGAGATACCCCTCAAGTATATCTGTTGCTTAGGTTGTCCTTCGCT
ATAACAGCCTCAAAAACCTTAAGTATAATCAATTAAAGGGCTAACCCAGCTCACCTGGCTATAC
CTTGACCATAACCATATCAGCAATTAGGCAAATGCTTTAATGAAATACGCAGACTCAAAGA
GCTGATTCTTAGTCCAATAGAATCTCTATTCTTAACAAATACCTCAGACCTGTGACAAATT
TACGGAACCTGGATCTGCCTATAATCAGCTGCATTCTGGATCTGAACAGTTGGGGCTTG
CGGAAGCTGCTGAGTTACATTACGGTCTAACCTCCCTGAGAACCATCCCTGTGCGAATATTCA
AGACTGCCAACCTGGAACCTGGACCTGGGATATAACCGGATCCGAAGTTAGCCAGGAATG
TCTTGCTGGCATGATCAGACTCAAAGAACCTCACCTGGAGCACAATTCAATTTCAGCTCAAC
CTGGCCCTTTTCAAGGTTGGCTAGCCTCAGAACCTTACTTGCGATGGAATAAAATCAGTGT
CATAGGACAGACCATGTCCTGGACCTGGAGCTCCCTACAAAGGCTGATTTATCAGGCAATGAGA
TCGAAGCTTCAGTGGACCCAGTGTGTTCCAGTGTGTCCGAATCGAGCCTCAACCTGGAT
TCCAACAAGCTCACATTATTGGTCAAGAGATTGGATTCTGGATATCCCTCAATGACATCAG
TCTTGCTGGGAATATATGGGAATGCAAGCAGAAATATTGCTCTTGTAACTGGCTGAAAGTT
TTAAAGGTCAAGGGAGAATACAATTACTGTGCCCAGTCAAGAGCTCAAGGAGATAATGTG
ATCGATGCAGTGAAGAACTACAGCATCTGGGAAAAGTACTACAGAGAGTTGATCTGGCAG
GGCTCTCCCAAAGCCGACGTTAACGACGCTCCCCAGGCCAGCATGAGACAAACCCCCCTT
TGCCCCCGACGGTGGAGCCACAGAGCCGCCAGAGACCGATGCTGACGCCGAGCACATCT
TTCCATAAAATCATCGCGGGCAGCGTGGCCTTCTGCTCGTCATCTGCTGGTTAT
CTACGTGTATGGAAGCGGTACCCCTGCGAGCATGAAGCAGCTGCAAGCGCTCCCTATGCGAA
GGCACAGGAAAAGAAAAGACAGTCCCTAACGAAATGACTCCAGCACCCAGGAATTATGTA
GATTATAAACCCACCAACACGGAGACCGAGCGAGATGCTGCTGAATGGGACGGACCTGCACCTA
TAACAAATCGGGCTCCAGGGAGTGTGAGGTATGAACCATTGTGATAAAAAGAGCTTAAAGCT
GGGAAATAAGTGGTCTTATTGAACTCTGGTACTATCAAGGGAACCGCATGCCCTCCCC
TTCCCTCTCCCTCACTTGGCAAGATCCTCCTGTCCGTTTACTGCATTCAATAACT
GGTCATTTCCTCTCATACATAATCAACCCATTGAAATTAAATACACAAATCAATGTGAAGCTT
GAACCTCGGTTAATATAACCTATIGTATAAGACCCCTTACTGATTCCATTATGTCGCATT
GTTTAAGATAAAACTTCTTCATAGGTAaaaaaaaaaaaaaa

FIGURE 124

MGFNVIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGCRCEGKMYCESQKLQEIPSSISAGCLG
LSLRYNSLQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYFLNNNTFR
PVTNLRNLDLDSYNQLHSLGSEQFRGLRKLLSLHRSNSLRTIPVRIFQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKEHLLEHNQFSKLNLAFFPRLVSLQONLYLQWNKISVIGQTMSWTWSSLQRLDL
SGNEIEAFSGPSVFQCVPNLQRNLNDSNKLTFIGQEILDWSWISLNDISLAGNIWECSRNICSLVN
WLKSFKGLRENTIICASPTELQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFHKIAGSVALFLSVLVILLVIYVSWKRYPASMKQLQQR
SLMRRHRKKKRQSLKQMTPSTQE FYVDYKPTNTETSEMLLNNGTGPCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTCGCTACTGCTGAATGTCCGTCCGGAGGAGGGAGGGCTTTGCCGCTG
ACCCAGAGATGGCCCCGAGCGAGCAAATTCCACTGTCCGGCTGCCGGCTACCGTGGCCGAGCT
AGCAACCTTCCCCTGGATCTCACAAAACTCGACTCCAATGCAAGGAGAACGAGCTTGC
GGTTGGGAGACGGTGCAAGAGAACTGCCCCCTATAGGGAAATGGTGCGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTCTAAAGCTTGGCAAGGAGTGACACCCGCCATTACAGACACGTAGT
GTATTCTGGAGGTGCAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTGGCAAAGTGAAG
ATGAGCATTATCCCCTTGGAAATCAGTCATTGGAGGGGATGGCTGGTGTATTGGCCAGTT
TTAGCCAATCCAACTGACCTAGTGAAGGTTCAGATGCAAATGGAAGGAAAAGGAACTGGAGG
AAAACCATTGCGATTCGTGGTTACATCATGCATTGCAAAATCTAGCTGAAGGAGGAATAC
GAGGGCTTGGCAGGCTGGTACCCATATACAAAGAGCAGCACTGGTGAATATGGGAGATTA
ACCACTTATGATACAAGTAAACTACTTGGTTATTGAATACACCACTGAGGACAATATCATGAC
TCACGGTTTATCAAGTTTATGTTCTGGACTGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCA
TCAAAAAGCAGAATAATGAATCAACCACGAGATAAACAAGGAAGGGGACTTTGTATAAAATCATCG
ACTGACTGCTTGATTCAGGCTGTTCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTACC
ATCTTGGCTGAGAATGACCCCTTGGTCAATGGTGTTCTGGTTACTTATGAAAAATCAGAGAGA
TGAGTGGAGTCAGTCCATTTTAA

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPPLDLTKTRLQMGEAALARLGDGARES
APYRGMVRTALGIIEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLREVFGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMQUEGKRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTTYDTVKHLYLVNTPLEDNIMTHGLSSLCSGLVASILGTPADVIKSRMNQP
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCCAAAGCAGGTGGCGGCAGGAGAGCGGCCGGCGTCAGCTCCTCGAC
CCCCGTGTGGGCTAGTCCAGCGAGGCGACGGCGGGCGTGGGCCCATGGCCAGGCCCGGCATGG
AGCGGTGGCGCGACCAGGCTGGCGCTGGTGAACGGGGCCTCGGGGGCATGGCGGGCGTGGCC
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGCTGGCGCCACTGTGGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTGCAAGGCTACCCGGGACTTGATCCCTACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGAGACATC
TGCATCAACAATGCTGGCTTGGCCGGCCTGACACCCCTGCTCTCAGGCAGCACCAGTGGTGGAA
GGACATGTTCAATGTAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA
AGGAGCAGGAATGTGGACGATGGGACATCATTAACATCAATAGCATGCTGGCCACCGAGTGT
CCCCCTGCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCAGCGCTGACAGAGGGACT
GAGGCAAGAGCTCGGGAGGCCAGACCCACATCCGAGGCCACGTGCACTCTCCAGGTGTGG
AGACACAATTGCCCTCAAACCTCCAGACAAGGACCCCTGAGAAGGCAGCTGCCACCTATGAGCAA
ATGAAGTGTCTCAAACCCGAGGAATGTGGCCAGGGCTTATCTACGTCCCTCAGCACCCCCG
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTAGTGACTGTGGGAGCTCC
TCCTTCCCTCCCCACCCCTCATGGCTTGCCTGCCTCTGGATTTAGGTGTTGAGATTTCTGGAT
CACGGGATACCACTTCCTGTCCACACCCGACCAGGGCTAGAAAATTGTTGAGATTTATA
TCATCTTGTCAAATTGCTTCAGTTGTAATGTAAAAATGGGCTGGGAAAGGAGGTGGTGTCCC
TAATTGTTTACTTGTAACTTGTCTTGTGCCCTGGCACTTGGCTTGTCTGCTCTCAGTG
TCTTCCCTTGACATGGGAAAGGAGTTGTGCCAAATCCCATCTTGCACCTAACGTCTG
TGGCTCAGGGCTGGGTGGCAGAGGGAGGCCTCACCTATATCTGTGTTATCCAGGGCTCC
AGACTTCCCTCTGCCTGCCCACTGCACCCCTCTCCCTTATCTATCTCCTCTCGGCTCCC
AGCCCAAGTCTTGGCTTCTGTCCCCTCTGGGTCACTCCACTCTGACTCTGACTATGGCAG
CAGAACACCAGGGCTGGCCAGTGGATTTCATGGTGATCATTAAAAAGAAAATCGCAACCAA
AAAAAAAAAA

FIGURE 128

MARPGMERWRDRRLALVTGASGGIGAAVARALVQQGLKVVGCARTVGНИЕELAAECKSAGYPGTЛИ
PYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLARPDTLLSGSTSGWKDMFNVLALSICTR
EAYQSMKERNVDDGHI ININSMGHRVLPLSVTHFYSATKYAVTALTEGLRQEЛREAQTHIRATC
ISPGVVETQFAFKLHDKPEKAATYEQMКCLKPEDVAЕAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACTTCTACATGGGCCTCCTGCTGGTGCCTTCCTCAGCCTGCGGTGGCTACACCAT
CATGTCCCTCCCACCCCTCCTTGACTGCGGGCGTCAGGTGCAGAGTCTCAGTGGCCGGGAGC
ACCTCCCTCCCGAGGCAGTCTGCTCACAGGGCCTCGGCCAGAATTCCAGTTGGTTCATGC
CAGCCTGAAAAGGCCATGGAACTTGGGTGAATCACCGATGCCATTAAAGAGGTTTCTGCCA
GGATGGAAATGTTAGGTCGTTCTGTCTGCCTGTTCATTCAGAGCCACCCAGCACGTGG
CCGTTGAGTGCTTGAATGAGGAACTGAGAAAATTAATTCTCATGTATTTCTCATTTA
TTAATTTTAACTGATAGGTACATATTGGGGTACATGTATTTGGATACATGTATACAA
TATATAATGATCAATCAGGGTAACTGGGATTATCCATCACATCAAACATTTTTTATTCTT
TTAGACAGAGTCTCACTGTACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC
CTCTGCCTGCCAGGTCAAGCGATTCTCATGCCTCACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC
GCACCACAATGCCCAACTAATTTGTATTTTAGTAGAGACGGGTTTGCCATGTTGCCAGG
CTGGCCTTGAACTCCTGGCCTCAAACAATCCACTTGCCCTGGCCTCCAAAGTGTATGATTACA
GGCGTGAGCCACCGTGCCTGGCCTAAACATTTATCTTCTTGTTGGAACTTGAATTAT
ACAATGAATTATTGTTACTGTCATCTCCCTGCTGCTATGGAACACTGGACTCTCCCT
ATCTAACTGTATTTGTACAGTTAACCACCGTACTTCATCCCCACTCCCTCTATCCTCCC
AACCTCTGATCACCTATTCTACTCTTACCTGAGATCCACTTTTTAGCTCCCACATGTG
AGTAAGAAAATGCAATTTGTTCTCTGTCGCCTGGTTATTCACTTAACATGACTCCTG
TTCCATCCATGTTGCTGCAAATGACAGGATTCGTTCTTAATTCAATTAAAAACCACACATG
GCAAAAA

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSPRGSSLRGPRPRIPVLVSCQPV
KGHGTLGESPMFKRVFCQDGTVRSFCVCAVFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTGTATAAAGACCTAACACTGCTGACCATGATCAGCCAGCCTGGAGC
ATCTTCCTCATCGGGACTAAAATTGGGCTGTCCTCAAGTAGCACCTCTATCAGTTAGGCTAAATCCTG
TCCATCTGTGTCGCTGCATGCGGGTTCTTACTGTAATGATCGCTTCTGACATCCATTCAAACAG
GAATACCAAGAGGATGCTACAACTCTACCTCAGAACAAATAAATGCTGGGATTCCCTCAGAT
TTGAAAAACTGCTGAAAGTAGAAAGAATAACCTATACCAACAGTTAGATGAATTTCCTACCAACCT
CCCCAAAGTATGTAAGAGTTACATTGCAAGAAATAACATAAGGACTATCACTTATGATTCACTTC
AAATTCCCTATCTGGAAGAATTACATTTAGATGACAACACTCTGCTCTGCAGTTAGCATAGAAGAGGGAGCA
TTCCGAGACAGCAACTATCTCGACTGCTTTCTGTCCCGTAATCACCTTAGCACAATTCCCTGGGGTT
GCCAGGACTATAGAAGAACTACGCTGGATGATAATCGCATATCCACTATTCATCACCATCTTCAAG
GTCTCACTAGTCTAAAACGCCCTGGTTCTAGATGGAACACATGAGTTAGGTGACAAAGTT
TTCTTCACCTAGTTAATTGACAGAGCTGCTCCCTGGTGCAGAAATTCCCTGACTGCTGCCAGTAAACCT
TCCAGGCACAAACCTGAGGAAGCTTATCTCAAGATAACCACATCAATCGGTGCCCAAATGCTTTT
CTTATCTAAGGCAGCTATCGACTGGATATGCTCAATAAACCTAAGTAATTACCTCAGGGTATCTT
GATGATTGGACAATATAACACAACGTGATTCTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATG
GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGTGGCTATGTGCCAAGCCCCAGAAA
AGGTTCTGGGATGGCTATTAAGGATCTCAATGCAAGAACTGTTGATTGTAAGGACAGTGGGATTGTAAGC
ACCATTCAAGATAACCAACTGCAATAACCAACACAGTGTATCCTGCCAAGGACAGTGGCAGCTCCAGTGAC
CAAACAGCCAGATATTAAGAACCCAAAGCTCAACTAAGGATCAACAAACACAGGGAGTCCTCAAGAAAAA
CAATTACAATTACTGTGAAGTCTGTCACCTCTGATACCATTCTATCTTGGAAACTTGCTCTACCTATG
ACTGCTTGAGACTCAGCTGGCTTAAACTGGCCATAGCCCGCATTTGGATCTATAACAGAAACAATTGT
AACAGGGAAACGCAGTGAGTACTTGGTCACAGCCCTGGAGCCTGATTCAACCTATAAAGTATGCATGGTC
CCATGGAAACCGAACCTCTACCTATTGATGAAACTCCTGTTGATTGAGACTGAAACTGCACCCCT
CGAATGTACAACCCCTACAACCACCCCTAACCGAGAGCAAGAGAAAGAACCTTACAAAAACCCAATTAC
TTGGCTGCCATCTGGTGGGCTGTGGCCCTGGTTACCATGCCCTCTGCTTGTGTTAGTGTGTTGATG
TTCATAGGAATGGATCGCTCTCAAGGAACCTGTGCAATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
GCAGAAGCTGGCACTAAGAAGGACAACCTATCCTGGAAATCAGGGAAACTCTTCAAGATGTTACCAAT
AAGCAATGAACCCATCTGAAGGAGGAGTTGTAATACACACCATATTCCTCTTAATGGAATGAATCTGT
ACAAAAACAATCACAGTGAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTGGTTAAACCTAAGGGAGGTGATG
GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRCDAFGIYCNDRFLTSIPTGIPEDATTLYL
QNNQINNAGIPSIDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTITYDSLSKIPYL
EELHLDNSVSAVSIEEGAFRDSNYLRLFLSRNHLSLTIWPGLPRTIEELRLDDNRISTISSPL
QGLTSLKRLVLDGNLNNHGLDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHIN
RVPPNAFSYLRQLYRLDMSNNNLSNLPQGI FDDLDNITQLILRNNPWYCGCKMKWVDWLQSLPV
KVNVRGIMCQAPEKVRGMAIKDLNAELFDCKDGSIVSTIQTITTAIPNTVYPAQGQWPAPVTQPD
IKNPKLTKDQQTTGSPSRKTITITVKSVTSDTIHISWKLALPMTALRLSWLKLGHSPAFGSITET
IVTGERSEYLVTALEPDSPYKVCVPMETSNLYLFDETPVCIETETAPLRMYNPTTLNREQEKE
PYKNPNLPLAAIIGGAVALVTIALLALVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKDNS
ILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNLYKNNHSESSSNRSYRDSGIPDSDHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCACTCCCCCTGCAGCCACCCCTCCCAGAGTCCTTGCAGGCCACCCAGGCTCTGGCA
GCCCTGCCGGGCCACTTGTCTTCCATGTCTGCCAGGGGGAGGTGGGAAGGAGGTGGGAGGGAGGGCG
TGCAGAGGCAGTCTGGGCTTGGCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGAGCAGAG
GCCGGCCATGGCCAGCCTGGGGCTGCTGCTCTGCTCTTACTGACAGCACTGCCACCGCTGTGGT
CCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT
GCGCTGGAGAGAGCCACCGTCTCCTAGAACAGAGGCTGCCGTGAAATCAACCTGGATGGCATGGT
GGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAGTGTCCGGAGAAGTGGGGCCAGGAGCCCCTGC
TGCAGCCGCTGAGCCTGCGCGTGGGATGCTGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGATCCAAAGTACCTAACAGAGGTTCCAGCTGACCCCTCAGCCGG
GTTTTGGAAGCTCCCACATGCCCTGGATCCACACTGATGCCCTTGGTGTACCCACGTTGGC
CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCCTGGTGCAGCTGCTGGAACCGGGACG
GACAGCAGCGAGCCCTGCGGCCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCGGCTGCTC
AGGCTACTGCCGTCCCACCAACTGCTCTTCTCTGGGCCAGAATGAGGGGATGCACACAGG
GACCACTCCAACAGAGCCAGGACTATACTAACCTCTGCGCCAACATGATGGACTTGAACCGC
AGAGCTGAGGCCATCGGATACGCCAACCCCTACCCGGACATCTCATGGAAAACATCATGTTCTG
TGGAAATGGCGGCTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
AACAGCAGGAAGGGATGCTTCGGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
TATCAGCAGCATTTTCGAGGAGAGTGAAAGAGGGAGAAAAAAACATTTCAGATTCTCGCTCTG
TGCTCAGGCTGGAGTACAGTGGCGAACATCTGGCTCACTGCAACCTTGCCTCTGGTTCAAGC
AATTCTCTGCCCTCATCCTCCCGAGTAGCTGGACTACAGGAGCGTGCCACCATACTGGCTAAT
TTTTATTTTTAGTAGAGACAGGGTTCATCATGTTGCTCATGCTGGTCTCGAACCTCTGAT
CTCAAGAGATCCGCCAACCTCAGGCTCCAAAGTGTGGGATTTAGGTGTGAGCCACCGTGTCTG
GCTGAAAAGCACTTCAAAGAGACTGTGTTGAATAAGGGCAAGGTTCTGCCACCCAGCACTC
ATGGGGCTCTCTCCCTAGATGGCTGCTCCCTCCACAAACACAGCCACAGCAGTGGCAGCCCTGG
GTGGCTTCTATACATCCTGGCAGAACCCCCCAGCAAACAGAGAGGCCACACCCATCCACACCG
CCACCAAGCAGCGCTGAGACGGACGGTCCATGCCAGCTGCCCTGGAGGAGGAACAGACCC
TTTAGTCCTCATCCCTAGATCCTGGAGGGCACGGATCACATCCTGGAGAAGGCATCTGGAGG
ATAAGCAAAGCCACCCGACACCAATCTTGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGG
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAGTCAACTGCAACTGAAAAA

FIGURE 134

MSARGRWEGGGRACRGSLGLARAQGAERVTSSERQPAMASLGLLLLLTLAPPLWSSSLPGLD
TAESKATIADLILSALERATVFLERQLPEINLDGMGVVRVLEEQLKSVREKWAQEPLLQPLSLRV
GMLGEKLEAAIQRSLHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGTDssePCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQOSQD
YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGSDFYKLRWLEAILSWQKQQEGCFG
EPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFQFSCLILP
SSWDYRSVPYLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTCAGAGCTGCTGTCATGGCGGCCCTGTGGGGCTTCTTCCCCTGCTGCTG
CTGCTGCTATCGGGGATGTCCAGAGCTCGGAGGTGCCGGGCTGCTGCTGAGGGATCGGAGG
GAGTGGGGCGCATAGGAGATCGCTCAAGATTGAGGGCGTGCAAGTTCCAGGGGTGAAGC
CTCAGGACTGGATCTGGCGCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTTCTTAAG
ACAGATGGAGTTTGTGGTCATGATATACTTCTGGATCTTATGTAGTGGAAAGTTGTATCTCC
AGCTTACAGATTGATCCCGTTCGAGTGGATATCACTCGAAAGGAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCAAATGAAATCTCAGGT
CCACCTCTTACTTTATTAAAAGGAATCGTGGGCTGGACAGACTTCTAATGAACCCAATGGT
TATGATGATGGTTCTCCTTATTGATATTGTGCTCTGCCTAAAGGGTCAACACAAGTGATC
CTGACATGAGACGGGAAATGGAGCAGTCATGAATATGCTGAATTCCAACCATGAGTTGCCTGAT
GTTTCTGAGTTCATGACAAGACTCTCTTCAAAATCATCTGGCAAATCTAGCAGCGGCAGCAG
TAAAACAGGCAAAGTGGGCTGGCAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTGCAC
AAACACGGCAACACTGGTGGCATCCAAGTCTTGGAAAACCGTGTGAAGCAACTACTATAAACTT
GAGTCATCCCGACGTTGATCTTACAACGTGTATGTT
AACTTTTAGCACATGTTGTACTGGTACACGAGAAAACCCAGCTTCATCTTGTCTGTAT
GAGGTCAATATTGATGTCACTGAATTAATTACAGTGTCTATAGAAAATGCCATTAATAAATTAT
ATGAACTACTATACATTATGTATATTAATTAAAACATCTTAATCCAGAAATCAAAAAAAA
AAAAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLSDVQSSEVPGAAAEGSGSGVGIGDRFKIEGRAVPGVKPQDWISAA
RVLVDGEEHVGFLKTDGSFVVHDIPSGSYVVEVSPAYRFDPVRVDITSKGKMRARYVNYIKTSE
VVRLPYPLQMKGSSGPPSYFIKRESWGWTDFLMNPMMVLPLLIFVLLPKVVNTSDPDMRREME
QSMNMLNSNHELPDVSEFMTRLFSSKSSGKSSGSSKTGSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTTCTGTGAGAATTGATTTCTCCCAGTCCCCGTGGGTCTGAGGGGA
CCAGAAGGGTGAGCTACGTTGGCTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCAAAACAA
GTTTGACATTTCCCCTGAAATGTCATTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC
CTTACCTGCTGGGACTAACGGGGAGGCCAGGATGGGGACAGAATAAAGGAGGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACCTCAGACCTGAAATCTCTTCAACGGGAGGCTTGGCAGT
TTTCTTACTCCTGTGGTCTCCAGATTCAGGCCTAAGATGAAAGCCTCTAGTCTGCCTTCAGC
CTTCTCTGCTGCGTTTATCTCTATGGACTCCTCACTGGACTGAAGACACTCAATTGGG
AAGCTGTGTGATGCCAACAAACCTTCAGGAAATACGAAATGGATTTCTGAGATACGGGGCAGTG
TGCAAGCCAAGATGAAACATTGACATCAGAATCTTAAGGAGGACTGAGTCTTGCAAGACACA
AAGCCTGCAATCGATGCTGCCCTGCCATTGCTAAGACTCTATCTGGACAGGGTATTAA
AAACTACCAGACCCCTGACCATTATACTCTCCGAAGATCAGCAGCCTGCCAATTCTTCTTA
CCATCAAGAAGGACCTCCGGCTCTCATGCCACATGACATGCCATTGTGGGGAGGAAGCAATG
AAGAAATACAGCCAGATTCTGAGTCACTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
TTTGGGGAACTAGACATTCTCTGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGTGCTG
CTGCTAAGAATATTGGAGGTCAAGAGCTCCAGTCTCAATACCTGCAGAGGAGGCATGCCCAA
ACCACCATCTTTACTGTACTAGTCTGTGCTGGTCACAGTGTATCTTATTGCATTACTG
CTTCCTTGCATGATTGTCTTATGCATCCCCATCTTAATTGAGACCATACTTGTATAAGATT
TGTAATATCTTCTGCTATTGGATATAATTATTAGTTAATATATTATTATTGCTATT
ATGTATTATTTTTACTGGACATGAAACTTTAAAAAAATCAGATTATTTATAACCTG
ACTAGAGCAGGTGATGTATTACAGTAAAAAAAAACCTGTAAATTCTAGAAGAGTGG
CTAGGGGGTTATTCAACTAAGGACATATTACTCATGCTGATGCTCTGTGAGAT
ATTTGAAATTGAACCAATGACTACTTAGGATGGGTGTGGAATAAGTTGATGTGGAATTGCAC
ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGATCTTCAG
CCAGGAATCCTACACGGCCAGCATGTATTCTACAAATAAAGTTCTTGCAACCAAAAAAA
AAAAAAAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPWGQNKGATTCATNSHSDSELRPEIF
SSREAWQFFLLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQCTK PANRCCLLRHLLRLYLD RVFKNYQTPDHYTLRKIS
SLANSFLTIKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGTCGGTTCCGCATCCAGCC
TAGCGTGTCCACGATCGGCTGGGACTTTCGCTACCTGTTGCGTAGCGATCGAGGTGC
TAGGGATCGCGTCTCCTCGGGATTCTCCCGCTCCGTTCGTCCCTGCCAGAGCGGAA
CACGGAGCGGAGCCCCCAGCGCCCGAACCTCGGCTGGAGCCAGTTCAACTGGACCACGCTGCC
ACCACCTCTCTCAGTAAAGTTATTGTTCTGATAGATGCCCTGAGAGATGATTGTTG
GGTCAAAGGGTGTGAAATTATGCCCTACACAACCTTACCTTGAGAAAAGGAGCATCTCACAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCCTGAATCAAGGCATTGATGACGGGGAG
CCTTCCCTGGCTTGTGCGACGTCAAGGAACCTCAATTCTCCGCTGACTGTTGAAAGACAGTGTGA
TAAGACAAGCAAAGCAGCTGGAAAAAGAATAGTCTTTATGGAGATGAAACCTGGGTTAAATTAA
TTCCCAAAGCATTGGAATATGAGAACACCTCATTTCTGTCAGATTACACAGAGGT
GGATAATAATGTCACGGGATTGGATAAAGTATTAAAAGAGGAGATTGGGACATATTAATCC
TCCACTACCTGGGGCTGGACACATGGCCACATTTCAGGGCACACAGCCCCCTGATTGGGAG
AAGCTGAGCGAGATGGACAGCGTGTGATGAAGATCCACACCTCACTGAGTCGAAGGAGAGA
GACGCCCTTACCCAATTGCTGGTCTTGTTGACCATGGCATGTCGAAACACAGAAGTCAG
GGGCCTCTCCACCGAGGAGGTGAATACACCTCTGATTTAATCAGTTCTGCGTTGAAAGGAAA
CCCAGGTGATATCGACATCCAAAGCACGTCCAATAGACGGATGTGGCTGGACACTGGCGATAGC
ACTTGGCTTACCGATCCAAAAGACAGTGTAGGGAGCCTCTATTCCAGTTGTGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTCACATTGAAATACAGTGCAGCTTAGTAAACTGTTGCAAGAG
AATGTGCGTCATATGAAAAAGATCCTGGGTTTGAGCAGTTAAATGTCAGAAAAGATTGCA
GAACGGATCAGACTGACTTGGAGAAAAGCATTAGAAGTCCTATTCAACCTGGCTCCAAGG
TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTGTCCTGAGTCACAAGTGGCCAG
TTCTCACCCCTGCTCTGCTCAGCGTCCCACAGGCACTGCACAGAAAAGGCTGAGCTGGAAGTCCA
CTGTCATCTCCTGGGTTCTGCTCTTTATTGGTGTGATCCTGGTCTTCGGCGTTACGT
CATTGTGTCACCTCAGCTGAAAGATTGCTGCTACTCTGTGGCCTCTGTTGGCTGGCGAGGCT
GCCTTCGTTACAGACTCTGGTGAACACCTGGTGTGCAAGTGTGGCAGTGCCCTGGAC
AGGGGGCCTCAGGGAAAGGACGTGGAGCAGCCTTATCCAGGCCTCTGGGTGTCCCACAGGTG
TTCACATCTGCTGTCAGGTCAAGTCAGATGCCCTAGTTCTGGAAAGCTAGGTTCTGCA
CAAGGTGATTGTAAGAGCTGGCGTCACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGAA
TCGGACAGCCTCCAGCAGAGGTGTGGAGCTGCAAGCTGAGGGAAAGAGAGACAATGGCCTGGA
CACTCAGGAGGGTCAAAGGAGACTTGTGTCACACTCATCCTGCCACCCCCAGAATGCATCCT
GCCTCATCAGGTCCAGATTCTTCAAGGGGACGTTCTGTTGGAATTCTAGTCCTGGCC
TCGGACACCTTCATTCGTTAGCTGGGAGTGGTGGTGAAGCAGTGAAGAAGAGGGGATGGTCAC
ACTCAGATCCACAGAGCCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGGCCCC
ACCCCAACCTGCAAGGCCCTCATCCCCCTCTGGCTTGAGCCGTCAAGGGCCCTGTGCTGAGTGT
CTGACCGAGACACTCACAGCTTGTGATCAGGGCACAGGCTTCTGGAGCCAGGATGATCTGTG
CCACGCTGCAACCTCGGGCCCATCTGGGCTCATGCTCTCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTAACAAAAGAATAAACGGCAATAATTGAGAAAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTTLPPPLF
SKVVIVLIDLALRDDFVFGSKGVKFMPYITTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDSVIRQAKAAGKRIVFYGDETWKLFPKHFVEYDGTTSFVSDYTEVDNNV
TRHLDKVILKRGDWDLILHYLGLDHIGHISGPNSPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGGTATCGTGACGCACCTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGTTGGAGTACA
ACT TCCCTATAGAAAACA
ACTGCCAGCACCTAAGACCAC
TCACACCTTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTCA
GCATTGACCA
GAGGATCACAAAGTACTGGTC
CTGGACTCTGGGAATCTCAT
AGCAGTTCCAGATA
AAAAACTACATACGCC
CAGAGATCTTCTTG
CATTAGCCTCATC
CTTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGTCTCTAAAGGGGAGTTTG
CTCTAC
TGTGACAAGGATA
AAAGGACAAAGTC
ATCCATCC
CTTCAGCTGAAGA
AGGAGAA
ACTGATGAAGCT
GGCTGCC
AAAAGGA
ATCAGCAC
GCCGG
CCCTTC
CATCTT
TATAGGG
CTCAGGT
GGGCT
CCTGG
AACATG
CTGGAG
TCGGCG
CTCAC
CCC
GGATGG
TCATCTG
CACCT
CTGCA
ATTGTA
ATGAGC
CT
GTTGGGGT
GACAGATA
AAATTG
GAGAAC
AGGAA
ACACATTG
AATTT
CATTC
AACCA
GAGCT
GAAATG
AGCCCC
AGTGAG
GTCAGC
GATTAGGAA
ACTGCC
CATTG
AACGC
CTTC
CTCG
CTA
TTTG
ACTA
ATTG
TATA
AAA
ACAC
CAA
ACCT
GCT
CACT

FIGURE 142

MLLLLLEYNFPPIENNQHLKTTHTFRVKNLNPKKFSIHDDQDHKVLVLDGNLIAVPDKNYIRPEI
FFALASSLSSASAEGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAQKESARRPFI
FYRAQVGWSWNMLESAAHPGWFIGTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36

N-myristoylation site.
amino acids 50-55, 87-92

Interleukin-1
amino acids 37-182

FIGURE 143

CTAGAGAGTATAGGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTTAA
TCCAGGATCCTGTCCTCCTGTCTGTAGGAGTGCCTGTTGCCAGTGTGGGTGAGACAAGTTG
TCCCACAGGGCTGTCTGAGCAGATAAGATTAAAGGGCTGGGTCTGTGCTCAATTAACTCCTGTGG
CACGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCATGCTGGGCCTGCCGTGG
AAGGGAGGTCTGTCTGGCGCTGCTGCTGCTCTTAGGCTCCCAGATCCTGCTGATCTATGC
CTGGCATTCCACGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTCGTTACCTCCCTGCCA
CAGTGGAGTTGCTGCCACACATTCAACCAACAGAGCAAGGACTACTATGCCAACAGACTGGGG
CACATCTTGAATTCTGGAAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGTGTGGAAATTGAAGACGACATTGACAACGCCATTCCAAGAAAGCACAG
AGCTGAACAATACTTCACCTGCTTCAACCATCAGCACCCAGGCCCCGGATGACTCAGTTCAGC
CTCCTGAACAAGACCTGCTGGAGGGATTCCACTGAGTGAAACCCACTCACAGGCTTGTCCATGT
GCTGCTCCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTCAGTGGCTGAGCAGCT
TTGGACTTGTGTTATCCTATTGATGTGTTGAGATCTCAGATCAGTGTAGTTAGAAAATCC
ACACATCTTGAGCCTAATCATGTAGTGTAGATCATTAAACATCAGCATTAAAGAAAAAAAAAAAA
AAA

FIGURE 144

MLGLPWKGGLSWALLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNNTFCFFTISTRP
WMTQFSLLNKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGCCAAGGTTCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGCTGGCACTGTCTGCATGCTGCTTCAGCCACCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAACAGTGGAACCGGAAGGCCCTGCCAGCACTGCCAGATCA
CTGAGGCCAGGTGGCTGAGAACGCCGGGAGCCTTCATCAAGCAAGGCCAAGCTCGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTGTCACCGGCTGCATCAATGCCA
CCCAGGCGCGAACCAAGGGGAGTTCCAGAACGCCAGAACAAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTGGAGAGGGCGCAGGACT
TCGGGTCAACCATGCACCAGCCAGTGCTCTGCCTCTGGCTTGATCTGGCTCATGGTGAAAT
AAGCTTGGCAGGAGGCTGGCAGTACAGAGCGCAGCAGCAGCAAATCTGGCAAGTGACCCAGCT
CTTCTCCCCAAACCCACGCGTGTCTGAAGGTGCCAGGAGCGCGATGCACTCGCACTGCAA
TGCCGCTCCACGTATGCGCCCTGGTATGTGCCTGCGTTCTGATAGATGGGGACTGTGGCTCT
CCGTCACTCCATTCTCAGCCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGATAAATGCTTGAT
GAGAAGAACACATCAGGCACTGCGCCACCTGCTCACAGTACTTCCAACAACACTTAGAGGTAG
GTGTATTCCCGTTTACAGATAAGGAAACTGAGGCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCCCTGGCTGTCAACCCAGGTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTGTACTTATGTCTCACAGGGACCCACATCCAAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTAGAAATAATGAAAACACCTGA

FIGURE 146

MRKHLSWWWLATVCMLLFSHLSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNKLHQO
VLWRLVQELCSLKHFCEFWLERGAGLRTMHQPVLCLLALIWLKV

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCAAAGGGCTGGGATTATAGGC GTGACCACCATGTC TGGTCCAGAGTCTCATT
CCTGATGATTTATAGACTCAAAGAAAACTCATGTTCAGAAGCTCTTCTCTGGCCTCCT
CTGTCTTCTTCCCTTTCTTATTAAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTCATTTGCTTGTCAGTGGGTAGGTCACTGAGTCTTAGTTTATTTTGAAATT
CAACTTCAGATTCAAGGGGTACATGTGAAGGTTGTTATGAGTATTGCATGATGCTGAGG
TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFLLILISSIYSESCKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTCACAGGAACCTCAGCACCCACAGGGCGGACAGCGCTCCCTCTACCTGGAGACTTGAC
TCCCGCGGCCAACCCCTGTTATCCCTTGACCGTCGAGTGTAGAGATCCTGCAGCCGCCAGTCC
CGGCCCCCTCTCCGCCAACACCCACCCCTCCTGGCTTCTCTGTTTACTCCTCTTTCAATTATA
ACAAAAGCTACAGCTCCAGGAGCCCAGCGCCGGCTGTGACCCAAGCGAGCGTGGAAAGAATGGGTT
CCTCGGGACCGGCACTGGATTCTGGTTAGTGCTCCGATTCAAGCTTCCCAAACCTGGAGGAA
GCCAAGACAAATCTACATAATAGAGAATTAAGTGAGAAAGACCTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAAAAACATATCCTCCAGAAAACAAGCCAGGTAGAGCAACTATTCTT
TGTTGATAACTTGAAACCTGCTAAAGGCAATAACAGAAAAGAAAAATTGAGAAAGAAAGACAATCTA
TAAGAAGCTCCCACCTGATAATAAGTGAAATGTGGAAAGATGTTGATTCAACCAAGAATCGAAAATG
ATCGATGATTATGACTCTACTAAGAGTGGATTGATCATAAATTCAAGATGATCCAGATGGCTCTCA
TCAACTAGACGGACTCCTTAACCCTGAAGACATTGTCATAAAATCGCTGCCAGGATTATGAAG
AAAATGACAGAGCCGTGTTGACAAGATTGTTCTAAACTACTTAATCTCGGCCTTACAGAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTACAAAATTAACTCAAGGAAGCCAAACAA
TTATGAGGAGGATCCAATAAGCCCACAGCTGGACTGAGAATCAGGCTGGAAAATACCAGAGAAG
TGACTCCAATGGCAGCAATTCAAGATGGCTTGCTAAGGGAGAAAACGATGAAACAGTATCTAACACA
TTAACCTGACAAATGGCTTGAAAGGAGAACTAAACCTACAGTGAAAGACAACCTTGAGGAACCTCA
ATATTCCCAAATTCTATGCGCTACTGAAAGTATTGATCAGAAAAGCAAGAGAAAAGAAA
CACTGATTACTATCATGAAACACTGATTGACTTGTGAAGATGATGGTGAATATGGAACAAATATCT
CCAGAAGAAGGTGTTCCCTACCTTGAAAACCTGGATGAAATGATTGCTTCAAGACCAAAACAGCT
AGAAAAAAATGCTACTGACAATATAAGCAAGCTTCCACGACCATCAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTTGAGGATTCCACAAAA
GATGATAACTCCAACCCAGGAGGAAAGACAGTGAAACCCAAAGGAAAACAGAAGCTTATTTGGAGC
CATCAGAAAAAAATATTGAATGGTTGAGAAACATGACAAAAGGGAAATAAGAAGATTATGACCTT
CAAAGATGAGAGACTTCATCAATAAACAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAA
GAAGCCGAGGCCATCAAGCGCATTATAGCAGCCTGTAAAATGGCAAAGATCCAGGAGTCTTCAA
CTGTTTCAGAAAACATAATATAGCTAAAACACTTCTAATTCTGTGATTAACCTTACAAGTGGTAAAACATAGCTTCTCC
GTTATTAGAAAAGTGTGAATTACAGTAGTTAACCTTACAAGTGGTAAAACATAGCTTCTCC
GTAAAAACTATCTGAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVLVLPIQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDFKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGKIPFKVTPMAAIQDGLAKGENDETWSNTLTNGLE
RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV
SYLENLDEMIALQTKNNKLEKNATDNISKLFPAPSEKSHEETDSTKEAAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKGKGNKEDYDLSKMRDFINKQADAYVEKGILDK
EEAEAIIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCGCCAGGAGAACATTCTGAGGGGAGTCTACACCCCTGTGGAGCTCAA
GATGGTCCTGAGTGGGGCGCTGTGCTTCGAATGAAGGACTCGGCATGAAGGTCTTTAATCTGC
ATAATAACCAGCTTCAGCTGGAGGGCTGCATGCAGGGAAAGGTCAATTAAAGGTGAAGAGATCAGC
GTGGTCCCCAATCGGTGGCTGGATGCCAGCCCTGTCCCCCTCATCCTGGGTGTCCAGGGTGGAG
CCAGTGCTGTCTGATGTGGGTGGGGCAGGAGCCACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTTGGTCCAAGGAATCCAAGAGCTTCACCTCTACCGGGGGCATGGGCTCACC
TCCAGCTTCGAGTCGGCTGCCAACCCGGCTGGTCTGTGCAAGGTGCTGAAGCCGATCAGCC
TGTCAGACTCACCCAGCTTCCCAGAATGGTGGCTGAATGCCCATCACAGACTTCTACTTCC
AGCAGTGTGACT**TAGGGCAACGTGCCCCCAGA**CTCCTGGGAGAGCCAGCTGGGTGAGGGT
GAGTGGAGGAGACCCATGGGGACAATCACTCTCTGCTCTCAGGACCCCCACGTGACTTAG
TGGGCACCTGACCACCTTGTCTCTGGTCTCCAGTTGGATAAATTCTGAGATTGGAGCTCAGT
CCACGGCTCCCCACTGGATGGTGCCTACTGCTGTGGAACCTTGTAAAAACCATGTGGGTAAA
CTGGGAATAACATGAAAAGATTCTGTGGGGTGGGGGGAGTGGTGGGAATCATTCTGCT
TAATGGTAACTGACAAGTGTACCCCTGAGCCCCGAGGCCAACCCATCCCCAGTTGAGCCTATA
GGGTCACTAGCTCCACATGAAGTCTGTCACTCACCACGTGCAAGGAGAGGGAGGTGGTCATA
GAGTCAGGGATCTATGGCCCTTGGCCCAGCCCCACCCCTTCAATCTGCCACTGTGATA
TGCTACCTTCTATCTCCCTCATCATTTGTGGCATGAGGAGGTGGTGTAGTCAGAA
GAAATGGCTGAGCTCAGAAGATAAAAAGATAAGTAGGGTATGCTGATCCTTTAAAAACCAA
GATACAATCAAATCCCAGATGCTGGTCTCATTTCCATGAAAAAGTGCTCATGACATATTGAGA
AGACCTACTTACAAAGTGGCATATAATTGCAATTATTAAATTAAAAGATAACCTATTATATT
TCTTATAGAAAAAGTCTGGAAGAGTTACTCAATTGTAATGCAATGTCAGGGTGGTGGCAGTAT
AGGTGATTTCTTTAATTCTGTTAATTCTGTATTCTCTAATTCTACAATGAAGATGA
ATTCTTGTATAAAAAGAAAATTAAATCTGAGGTAAAGCAGAGCAGACATCATCTCTGA
TTGTCCTCAGCCTCCACTTCCCAGAGTAATTCAAATTGAATCGAGCTCTGCTGCTCTGGTGG
TTGTAGTAGTGTAGCAGAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAGTTGT
GTGGCTGGAATCTCTGGTAAGGAACCTTAAAGAACAAAATCATCTGTAATTCTTCTAGAAG
GATCACAGCCCCGGATTCCAAGGCATTGATCCAGTCTAAGAAGGCTGCTGTACTGGTGTGA
ATTGTGCCCCCTCAAATTCACATCCTCTGGATCTCAGTGTGAGTTATTGGAGATAAG
GTCTCTGAGATGTAGTTAGTAAGACAAGTCATGCTGGATGAAGGTAGACCTAAATTCAATAT
GACTGGTTCTGTATGAAAAGGAGGAGCACAGAGACAGAGGAGACGCGGGGAAGACTATGTA
AAGATGAAGGCAGAGATCGGAGTTTGCAAGGACAAGCTAAGAACACCAAGGATTGGCAACC
ATCAGAAGCTTGGAGAGGCAAGAAGAATTCTCCCTAGAGGCTTAGAGGGATAACGGCTCTG
CTGAAACCTTAATCTCAGACTTCCAGCCTCTGAACGAAGAAAGAATAAATTCCGGCTTTAA
GCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAAACTAATACAGCTGCTAAAATGATCCCCTGT
CTCCTCGTGTTCATTCTGTGTGTCCTCCACAAATGTACCAAAGTTGTCTTGACCAA
TAGAATATGGCAGAAGTGTAGGGCATGCCACTTCAAGATTAGGTATAAAAAGACACTGCACTTC
TACTTGAGGCCCTCTCTGCCCCACCCACGGCCCCCAATCTATCTGGCTCACTCGCTCTGGGG
AAGCTAGCTGCCATGCTATGAGCAGGCCATAAAGAGACTTACGTGTTAAAATGAAGTCTCCT
GCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATCGATGTTGTTGTT
AAGTTGCTCAGTTGGCTAACCTGTTATGCAAGATAAAATAATGCAAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEIESVVPNRWLDASLSPVILGVQGGS
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFCTVPEADQP
VRLTQLPENGGWNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTTCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTCTTCCTTATGGGGACCTGGCCACCAAGCTGCCTCCTCTGG
CCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATCACCAACCGCACCTCATGCTGGCTAAGGGAGCTAGCTTGGCTGATAA
CAACACAGACGTTCGTCTCATTGGGGAGAAAAGTGTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACCTCACCCCTGAAGAAGTGCTGTTCCATCTGATAGGTT
CAGCCTTATATGCAGGGAGGTGGTGCCTTCCCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTTGCTGTTATGTCTTGAGAAAT
GCCTGCATTTGACCAGAGCAAAGCTGAAAATGAATAACTAACCCCCCTTCCCTGCTAGAAATAA
CAATTAGATGCCCAAAGCGATTTTTAACAAAAGGAAGATGGAGCCAAACTCCATCATG
ATGGGTGGATTCCAATGAACCCCTGCCTAGTTACAAAGGAAACCAATGCCACTTTGTTATA
AGACCAGAAGGTAGACTTCTAAGCATAGATATTATTGATAACATTTCATTGTAACTGGTGTTC
TATACACAGAAAACAATTATTTTAAATAATTGTCTTTCCATAAAAAAGATTACTTTCCAT
TCCTTAGGGGAAAAACCCCTAAATAGCTCATGTTCCATAATCAGTACTTATATTATAAA
TGTATTATTATTATAAGACTGCATTATTATCATTATTATAATATGGATTATTAT
AGAAACATCATTGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTATGACAATAATTAT
AGAGCTATAACATGTTATTGACCTCAATAAACACTGGATATCCC

FIGURE 154

MAALQKS VSS FLMGTLAT SCLLL ALLVQGGAA APISSH CRLDKS NFQQPY ITNRTF MLAKE ASL
ADNN TDVRLIGEKL FHGVSM SERCYLMK QV LNF TLEE VLFPQSDRF QPYM QEV VPFL ARLS NRLS
TCHIEGDDLH IQRNV QKL KDTVKKL GESGE IKAIGEL DLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT
CAGTCAGTGCCCGACTTGTGACTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGCAGAGGGC
TGCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC
TGCCAGGTTTGGGGCTGGGGGCCAAGTGGAGTGGAGAACTGGGATCCCAGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGAAGTTCTCTCATTAGCCTTCTACAGGTGGTTGCAT
TCTTGGCAATGGTCAATGGGAACCCACACCTACAGCCTGGCCAGCTGCTGCCCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCTGTGCTCCCTAGAGCCTGCTAG
GCCCAACCGCCACCCAGAGTCCTGTAGGCCAGTGAAGATGGACCCCTAACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCCAGGACCTGTACCACGCCGT
TGCCTGTGCCCGCACTGCGTCAGCCTACAGACAGGCTCCCACATGGACCCCGGGCAACTCGGA
GCTGCTCTACCACAACCAGACTGTCTTCTACAGGCGGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTTAGCTTGTGTGTGCGGCCCGT
GTGATGGGCTAGCGGACCTGCTGGAGGCTGGTCCCTTTGGAAACCTGGAGCCAGGTGTACA
ACCACTTGCCTGAAGGGCAGGATGCCAGATGCTTGGCCCTGTGAAGTGTGTGGAGCAG
CAGGATCCCGGGACAGGATGGGGCTTGGGAAAACCTGCACTTCTGCACATTGAAAAGAG
CAGCTGCTGCTTAGGGCCGCCGAAGCTGGTGTCTGTCACTTCTCAGGAAAGGTTTCAA
GTTCTGCCATTCTGGAGGCCACCACTCCTGTCTTCCCTCTTCCATCCCTGCTACCTG
GCCCACAGGCACAGGCACATTCTAGATAATTCCCCCTGCTGGAGAAGAAAGAGCCCTGGTTTATT
TGTTGTTACTCATCACTCAGTGAGCATCTACTTGGGTGCATTCTAGTGTAGTTACTAGTCTT
TTGACATGGATGATTCTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTATCAAATAAATAT
CTTTATTAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGTHTYSHWPSCCP SKQDTSEELL RWSTVPVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCG**A**TTCGCTCGTGTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCGAGAGCC
GACC GTTCAATGTGGCTCTGAAACTGGGCATCTCCAGAGTGGATGCTACAACATGATCTAA TCC
CCGGAGACTTGAGGGACCTCGAGTAGAACCTGTTACAACACTAGTGTGCAACAGGGACTATTCA
ATTTGATGAATGTAAGCTGGGTACTCCGGGAGATGCCAGCATCCGCTTGTGAAGGCCACCAA
GATTGTTGTGACGGGAAAAGCAACTTCCAGTCCTACAGCTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCCTCTGGTGGATAATGGACATTTCCTACATGGCTTCCCTGTA
GAGCTGAACACAGTCTATTICATTGGGCCATAATATTCTAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTTCACCTCACCGAGCTGCCTAGACCACATAATGAAATATAAAAAAA
AGTGTGTCAAGGCCGGAACGCTGTGGATCCGAACATCACTGCTTGTAAAGAAGAATGAGGAGACA
GTAGAAGTGAACCTCACAACCCTCCCGGAAACAGATACTGGCTTATCCAACACAGCAC
TATCATGGGTTTCTCAGGTGTTGAGCCACCCAGAAGAAACACGCGAGCTTCAGTGGTGA
TTCCAGTGA CTGGGATAGTGAAGGTGCTACGGTGAGCTGACTCCATATTTCCTACTTGTGGC
AGCGACTGCATCCGACATAAGAACAGTTGTGCTCTGCCACAAACAGGCGTCCCTTCCCTCT
GGATAACAAACAAAGCAAGCCGGAGGCTGGCTGCCTCCTCCTGCTGTCTGCTGGTGGCCA
CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTCCCTT
TCTACCACCACTACTGCCCCCCATTAAGGTTCTGTGGTTACCCATCTGAAATATGTTCCA
TCACACAAATTGTTACTTCACTGAATTCTCAAAACATTGCAGAAGTGAGGTGATCCTGAA
AGTGGCAGAAAAGAAAATAGCAGAGATGGTCCAGTGCAGTGGCTGCCACTCAAAGAAGGCA
GCAGACAAAGCTGCTTCCCTTCCAATGACGTCAACAGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCCAGTGAGAACTCTCAAGACCTTCCCCCTGCTTAAACCTTTCTGCA
GTGATCTAAGAACGCCAGATTCACTGCACAAATACGTGGTGGTCACTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTCAGTGTCTGCCCAAGTACCAACCTCATGAAGGATGCCACTGCTT
CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAGATCACAAGCCTGCCACG
ATGGCTGCTGCTCCTTGTAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLIPGDLRDLRVEPVTTSVATGDYSILMNVS梧
LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP
NANMNEGDPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKNEETVEVNFTTPLGNRYMALIQH
STIIGFSQVFEPHQKKQTRASVVIPTVGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPQTVPPFLDNNK
SKPGGWLPLLLLSLLVATWVLVAGIYLMWRHERIKTSFSTTLLPPIKVLVVYPSEICFHHTICYFTEFL
QNHCRSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLSNDVNSVCDGTCGKSEGSPSENSQDLFPLA
FNLFCSIDLRSQIHLHKYVVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAEELLHVVKQQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-492

FIGURE 159

AGCCACCAGCGAACATGACAGTGAAGACCCCTGCATGGCCCAGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGGCTTGCCTTCTGAGTGAGGGCGCAGCTCGAATCCCCAAAGTAGGACATA
CTTTTTCCAAAAGCCTGAGAGTTGCCCGCTGTGCCAGGAGGTAGTATGAAGCTGACATTGGC
ATCATCAATGAAAACCAGCGCTTCCATGTCACGTAACATCGAGAGCCGCTCCACCTCCCCCTG
GAATTACACTGTCACTTGGGACCCCAACCGTACCCCTCGGAAGTTGTACAGGCCAGTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCATCCAGCAA
GAGACCCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTCTTCCAGTTGGAGAAGGTGCT
GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAGTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 16o

MTVKTLHGPAMVKYLLLSILGLAFLSEAARKIPKVGH~~FFQK~~PESCPVPGGSMKLDIGIINEN
QRVSMSRNIESRSTSPWNYTVTWD~~PNRY~~PSEVVQAQCRLGCINAQGKEDISMNSVPIQQETLVV
RRKHQGCSVSFQLEKVLVTVGCTCVTPVIHHQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAGGACTCCCAGG
ACAGAGAGTGACAAACTACCCAGCACAGCCCCCTCCGCCCTCTGGAGGCTGAAGAGGGATTC
CAGCCCCCTGCCACCCACAGACACAGGGTGAAGTGGCTGACTGGGGTGTCTGCCCTGGGGGGGGCAGCAC
AGGGCCTCAGGCCTGGGTGCCACCTGGCACCAGAAGATGCCTGTGCCCTGGTTCTGCTGTCC
TGGCACTGGGCCGAAGCCAGTGGCTTCTCTGGAGAGGCTTGTTGGGCCCTCAGGACGCTACC
CACTGCTCTCGGGGCTCTCTGCCCTCTGGGACAGTGACATACTCTGCCCTGCCCTGGGACAT
CGTGCCTGCTCCGGGCCCCGTGCTGGCCTACGCACCTGCAGACAGAGCTGGTGTGAGGTGCC
AGAAGGAGACCGACTGTGACCTCTGTGCTGGCTTGCCACTTGCCGTGCATGGCACTGG
GAAGAGCTGAAGATGAGGAAAAGTTGGAGGAGCAGCTGACTCAGGGGTGGAGGAGCCTAGGAA
TGCCCTCTCCAGGGCCAAGTCGTGCTCTCCAGGGCTACCCCTACTGCCCGCTGGGTCC
TGGAGGTGCAAGTGCCCTGCTGCCCTGTGCAAGTTGGTCACTGTGCCCCCTGTGGTATATGAC
TGCTCGAGGCTGCCCTAGGGAGTGAAGTACGAATCTGGCTCTACTCAGCCCAGGTACGAGAA
GGAACATCAACCACACACAGCAGCTGCCCTGCCCTGGCTCAACGTGTCAGCAGATGGTGACA
ACGTGCATCTGGTTCTGAATGTCCTGAGGAGCAGCAGCTGGCTCTCCCTGTACTGGAAATCAG
GTCCAGGGCCCCAAAACCCGGTGGCACAAAACCTGACTGGACCGCAGATCATTACCTTGAA
CCACACAGACCTGGTCCCTGCCCTGTATTCAAGGTGTCCTCTGAACCTGACTCCGTTAGGA
CGAACATCTGCCCTCAGGGAGGACCCCGCGCACACCAACCTGGCAAGCCGCCGACTG
CGACTGCTGACCCCTGAGAGCTGGCTGCTGAGCAGCAGCTGCTGCCCTGGCAGAAGCGGCACT
GTGCTGGCGGGCTCCGGTGGGACCCCTGCCAGCCACTGGTCCCACCGCTTCTGGAGAACG
TCACTGTGGACAAGGTTCTCGAGTCCATTGCTGAAGGCCACCCCTAACCTGTGTTAGGTG
AACAGCTCGGAGAACGCTGAGCTGAGGAGTGTGGCTGACTCCCTGGGCTCTCAAGAGA
CGATGTGCTACTGTTGGAGACACGAGGCCAGGACAACAGATCCCTGTGCTTGGAAACCA
GTGGCTGTACTTCACTACCCAGAAAGCCTCACGAGGGCAGCTGCCCTGGAGAGTACTTACTA
CAAGACCTGCACTGCCAGTGCTGCACTATGGACGATGACTTGGAGCGCTATGGCCTG
CCCCATGGACAATACATCCACAAGCGCTGGGCCCTCGTGTGGCTGCCCTGCCTACTCTTGCCG
CTGCGCTTCCCTCATCCTCTCAAAAGGATCACCGAAAGGGTGGCTGAGGCTTGTGAAA
CAGGACGTCGCTGGGGCGCCAGGGCCGCGCAGGGCCCTGCTCCCTACTCAGCCGATGA
CTCGGGTTTCGAGGCCCTGGTGGCGCCCTGGCTCGGCCCTGTGCCAGCTGCCCTGCGCTGG
CCGTAGACCTGTGGAGCCGTCGTGAACCTGAGCGCGCAGGGCCCTGGCTTACGCGCAG
CGGCCAGACCCCTGAGGAGGCCGCGTGGTCTTGTCTCTCCGGTGCCTGGCCT
GTGCAGGAGTGGTACAGGATGGGTGTCCGGGCCGGCGCACGGCCCGCACGACGCCCTCC
GCCCTCGCTCAGCTGCGTGCCTGCCGACTTCTGCAAGGGCCGGCCGGCAGCTACGTGGGG
GCCCTCGACAGGCTGCTCCACCCGAGCCGTACCCGCCCTTCCGCACCGTGCCGTCTT
CACACTGCCCTCCCAACTGCCAGACTTCTGGGCCCTGCACTGCCCTCGGCCCGCGTCCG
GGCGGCTCCAAGAGAGAGCGGAGCAAGTGTCCGGCCCTCAGCCAGCCCTGGATAGCTACTTC
CATCCCCGGGACTCCCGCGCCGGACGCCGGGTGGGACCGAGGGCGGGACCTGGGGGGGA
CGGGACTTAATAAAGGCAGACGCTGTTTCTAA

FIGURE 162

MPVPWFLLSLALGRSPVVLSLERLVPQDATHCSPGLSCLRWLSDS DILCLPGDIVPAPGPVLAPTHLQTELV
LRCQKETDCDLCLRVAVHLAVHGHWEEPEDEEKFGGAADSGVEEPRNASLQAQVVLSFQAYPTARCVILLE
QVPAALVQFGQSVGSVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPALPWLNVSADGDNVHLVLNVS
EEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQN
LWQAARLRLTLQSWL LDAPCSLPAAEALCWRA PGDPQCPLVPPLSWENVTVDKVLEFPPLLKGHPNLCVQ
VNSSEKLQLQECWLADSLGPLKDDVLLTRGPQDNRS LCALEPSGCTSLPSKASTRAARLGEYLLQDLQS
GQCLQLWDDDLGALWACPMDKYIHKRWALVWLACLLFAAALS LLLLKKDHAKGWLRLLKQDVRS GAAARG
RAALLLYSADD SGFERLV GALASALCQLPLRVA VDLWSRRELSAQGPVAWFHAQR RQTLQEGGVVVLLFSP
GAVALCSEWLQDG VSGPGAHGPHDAFRASI SCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTV PVFT
LPSQLPDFL GALQQP RAPRS GRLQ ERAEQ VSR ALQP ALDSYFHPPGT PAPGRGV GPGAG PGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTTGACTGTGGGATCCCTGGCT
GCTCACGCCCTGAGGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCCAGTCCAGCAACTTGA
AAACATCCTGACGTGGGACAGCGGGCAGAGGGCACCCAGACACGCTACAGCATCGAGTATA
AGACGTACGGAGAGAGGGACTGGGTGGCAAAGAAGGGCTGTCAGCGATCACCCGAAAGTCCTGC
AACCTGACGGTGGAGACGGCAACCTCACGGAGCTCTACTATGCCAGGGTCACCGCT
GTCAGTGCAGGGAGGGCGGTCAAGCCACCAAGATGACTGACAGGTCAGCTCTGCAGCACACTAC
CCTCAAGCCACCTGTATGTGACCTGTATCTCAAAGTGAAGATCGATTAGATGATTGTTCATCTA
CCCCCACGCCAACCTCGTGCAGGGCATGGCACCCGGCTAACCTGGAAAGACATCTCCATGACCTG
TTCTACCAACTTAGAGCTCAGGTCAACCGCACCTACCAAATGCACCTGGAGGGAAAGCAGAGAGA
ATATGAGTTCTCGGCTGACCCCTGACACAGAGTTCTGGCACCATCATGATTGCGTTCCA
CCTGGGCAAGGAGATGCCCCCTACATGTGCGAGTGAAGACACTGCCAGACCCGACATGGGACC
TACTCCTCTCGGAGCCTTCCATGGGCTTCCATCGTCGAGTACTCTGCTACCTGAG
CTACAGATACTGACCAAGCCGCTGCACCTCCAACTCCCTGAACGTCAGCGAGTCCCTGACTT
TCCAGCCGCTGCCTCATCCAGGAGCACGCTCTGATCCCTGTTGACCTCAGGGGGCCCCAGC
AGTCTGGCCAGCCTGTCCAGTACTCCCAGATCAGGGTGTCTGGACCCAGGGAGCCGAGGAGC
TCCACAGCGGATAGCCTGTCCAGATCACCTACTTAGGGCAGCCAGACATCTCCATCCTCAGC
CCTCAAACGTGCCACCTCCCCAGATCCTCTCCCACTGTCCCTATGCCCAAACGCTGCCCTGAG
GTCGGGGCCCCATCTATGCACCTCAGGTGACCCCGAAGCTCAATTCCATTCTACGCCCCACA
GCCATCTTAAGGTCCAGCCTCCTCTATGCCCTCAAGCCACTCCGGACAGCTGGCCTCCT
CCTATGGGTATGCAAGGGTCTGGAAAGACTCCCCACTGGACACTTTCTAGTCCTAAA
CACCTTAGGCCTAAAGGTCAAGGTCTCAGAAAGGCCACCAAGCTGGAAAGCTGCATGTTAGGTGGCCT
TTCTCTCAGGAGGTGACCTCTGGCTATGGAGGAATCCAAGAAGCAAATCTTGCACCAAGC
CCCTGGGATTTGACAGACAGAACATCTGACCCAAATGTGCTACACAGTGGGAGGAAGGGACA
CCACAGTACCTAAAGGGCCAGCTCCCCCTCCTCTCAGTCCAGATCGAGGGCCACCCCATGTC
CCTCCCTTGCAACCTCCTCCGGTCCATGTTCCCCCTGGACCAAGGTCCAAGTCCCTGGGGCC
TGCTGGAGTCCCTTGTTGTCAGGATGAAGCCAAGAGCCCAGCCCCCTGAGACCTCAGACCTG
GAGCAGCCCACAGAACTGGATTCTCTTCAGAGGCTGGCCCTGACTGTGCAGTGGAGTCCTG
AAGGGAAATGGGAAAGGCTTGGTGTCTCCCTGTCCTACCCAGTGTACATCCTGGCTGTCA
ATCCCATGCCCTGCCATGCCACACACTCTGCATCTGGCTCAGACGGGTGCCCTGAGAGAACG
AGAGGGAGTGGCATGCAGGGCCCTGCCATGGGTGCGCTCCTCACCGAACAAAGCAGCATGATA
AGGACTGCAGGGGGAGCTGGGGAGCAGCTTGAGACAAGCGCGTCTGAGCCCTG
CAAGGCAGAAATGACAGTGAAGGAGGAAATGCAGGGAAACTCCCGAGGTCCAGAGCCCCACCTC
CTAACACCATGGATTCAAAGTGTCAAGGGAAATTGCTCTCTGGCCCAATTCTGGCCAGTT
ACAATCTAGCTCGACAGAGCATGAGGCCCTGCCCTCTGTCAATTGTTCAAAGGTGGAGAGA
GCCCTGGAAAAGAACCCAGGGCTGGAAAAGAACCCAGAAGGGAGCTGGCAGAACCAACCTGC
ACTTCTGCCAAGGCCAGGGCAGGCCAGGACGGACTCTAGGGAGGGTGTGGCCTGAGCTCA
TTCCCAGCCAGGGCAACTGCCCTGACGTTGCACATTTCAGTCACTCCCTGTAGAACAAGC
GAAATGCAGGTCCACCAAGGGAGGAGACACACAAGCCTTCTGCAAGCAGGAGTTCAAGCCT
ATCCCTGAGAATGGGTTGAAAGGAAGGTGAGGGCTGTCCTGGCCCTGAGGGTACAATAACAC
TGTACTGATGTCAAACTTGCAGCTCTGCCCTGGGTCAAGCCATCTGGGCTCAAATTCCAGC
CTCACCAACTCAAAGCTGTGACTTCAAACAAATGAAATCAGTGCCAGAACCTCGGTTCTC
ATCTGTAATGTGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGTCATG
TCTTAAAGTCCTTAAAGTGCCTGGTACATGGCAGTGCCAATAACGGTAGCTATTAAAAA
AAAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGERDW
VAKKGCQIRITRKSCNLTVEGNLTELYYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVT CIS
KVRSIQMIVHPTPTPIRAGDGHLTLIEDIFHDLFYHLELQVNRTYQMHLGGKQREYEFFGLPDT
EFLGTIMICVPTWAKESAPYMCRVKTLPDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
PNSLNVQRVLTFQPLRFIQEHVILIPVFDLSPGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFPFYAPQAISKVQPSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLSVPKHLRPKGQLQKEPPAGSCMLGGSLQEVTSIAM
EESQEAKSLHQPLGICTDRTSDPNVLHSGEEGTPQYLKGQLPLLSSVQIEGHPMSLPLQPPSGPC
SPSDQGPSPWGLLESLVCPKDEAKSPAPETSDLEQPTELDSLFRGLALTVQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAGTCACCCGGGCCCGCGTGGCCACAACATGG
CTGCGGCACGGCTCTTGCTGGCTGGCTGGCTGGTGGGTCCCAGGGCAG
TCGGATCTCAGCCACGGACGGCGTTCTGACCTCAAAGTGTGCGGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGGAAAGCTTGAAGACTTCACGGCCCTGATTGTCGTTGTGAATTAA
AAAAAGGTGACGATGTATATGTTACTACAAACTGGCAGGGGATCCCTGAACCTTGGCTGGA
AGTGTGAACACAGTTGGATATTTCCAAAAGATTGATCAAGGTACTTCATAAAATACACGGA
AGAAGAGCTACATAATTCCAGCAGATGAGACAGACTTTGTCGTTGAAGGAGGAAGAGATGATT
TTAATAGTTATAATGTAGAACAGACTTTAGGATCTTGGAACTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAACAGTTCTCAGCACAGAGAAATCTCCTGAGGAGTCTGGGGCGTGA
ACTTGACCCCTGTGCCTGAGCCCGAGGCATTCAAGAGCTGATTCAAGAGGATGGAGAACGGTCTTCT
CAGAGAGCACCAGGGCTGCAGGGACAGCCTCAGCTCAGGAGAGCCACCCACACCAGCGGT
CCTGCGCTAACGCTCAGGGAGTGCAGTCTCGTTGGACACTTTGAAGAAATTCTGCACGATAA
ATTGAAAGTGCAGGAAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTGGAGCGGGAGA
AGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCCTTATT
CATTACAGCAAAGGATTCGTTGGCATCAAAATCTAAGTTGTTTACAAAGATGTTTAGTA
CTAAGCTGCCTGGCAGTTGCATTTGAGCCAACAAAATATTATTCCCTTAAGTA
AAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWLFVLGALWWVPGQSDLSHGRRFSDLKVC~~GDEEC~~MLMYRGKAEDFTGPDCRFVN
FKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEELHIPADETDFVC~~FEGGRD~~
DFNSYNVEELLGSLELEDVPEESKKAEEVSQHREKSPEESRGRELDPVPEPEAFRADSE~~DGE~~GA
FSESTEGLQGQPSAQESH~~P~~H~~T~~SGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGNSSPASVER
EKTDAYKVLKTEMSQRGSGQC~~V~~IHYSKGFRWHQNLSLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTAGAGGCCGGGAAGAGAAGCAAAGCGC
AACGGTGTGGTCCAAGCCGGGCTTCGCTCGCCTCTAGGACATACACGGGACCCCTAACTTC
AGTCCCCAAACGCGCACCCCGAAGTCAGTCTGAACCTCCAGCCCCGCACATCCACGCGCGCACAGG
CGCGGCAGGCAGGTCCCAGGCGAAGGCGATGCGCGCAGGGGTCGGGCAGCTGGCTCGGGC
GGCAGGGAGTAGGGCCCGCAGGGAGGCAGGGAGGCTGCATATTCAAGAGTCGCGGGCTCGCCCTG
GGCAGAGGCCGCCCGCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCGATGAGCCGCGTGG
TCTCGCTGCTGCTGGCGCCCGCTGCTGCGGCCACGGAGCCTCTGCCGCCGCGTGGTCAGC
GCCAAAAGGTGTGTTGCTGACTCAAGCATCCCTGCTACAAAATGCCACTTCCATGAAC
GTCCAGCCGAGTGAGCTTCAGGAGGCAGCCTGGCTGTGAGAGTGAGGGAGGAGTCCTCCTCA
GCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGAAAACCTGACAAAACCCGGG
ACAGGGATTCTGTGGTGGTATTCTGGATAGGGCTTGGAGGAATGGAGATGGCAAAACATCTGG
TGCCTGCCAGATCTTACCAAGTGGCTGATGGAAGCAATTCCAGTACCGAAACTGGTACACAG
ATGAACCTCCTGCGGAAGTGAAGAAGTGTGTTGATGTATCACCAACCAACTGCCAATCCTGGC
CTTGGGGTCCCTACCTTACCAAGTGGATGACAGGTGTAACATGAAGCACAATTATATTG
CAAGTATGAACCAGAGATTAATCCAACAGCCCTGTAGAAAAGCCTTATCTTACAAATCAACCAG
GAGACACCCATCAGAATGTGGTGTACTGAAGCAGGTATAATTCCAATCTAATTATGTTGTT
ATACCAACAATACCCCTGCTCTTACTGATACTGGTTGCTTGGAACCTGTTGTTCCAGATGCT
GCATAAAAGTAAAGGAAGAACAAAATAGTCAAACCAACTGCTACACTGTGGATTCAAAGAGTA
CCAGAAAAGAAAGTGGCATGGAAGTATAACTCATTGACTTGGTCCAGAATTGTAATTCT
GGATCTGTATAAGGAATGGCATCAGAACAAATAGCTTGGAAATGGCTGAAATCACAAAGGATCTGC
AAGATGAACTGTAAGCTCCCCCTTGAGGCAAATATTAAAGTAATTGTTATATGTCTATTATTC
TTTAAAGAATATGCTGTGCTAATAATGGAGTGGACATGCTTATTGCTAAAGGATGCACCCAA
ACTTCAAACCTCAAGCAAATGAAATGGACAATGCACTGAGATAAAAGTTGTTATCAACACGTCGGGAGTA
TGTGTGTTAGAAGCAATTCTTTTATTCTTACCTTCATAAGTTGTTATCTAGTCAATGTA
TGTATATTGTTAGAAGCAATTACAGTGTGCAAAGTATTACCTTGCATAAGTGTGATAAAA
ATGAACCTGTTCTAATATTATTATGGCATCTCATTTCATAACATGCTCTTGTGATAAAG
AAACTTATTACTGTTGTCACGTGAATTACACACACACAAATATAGTACCATAGAAAAGTTGT
TTCTCGAAATAATTCTCATCTTACAGCTCTGCTTGGTCAATGTCTAGGAAATCTCTCAGA
AATAAGAAGCTATTCTATTGCAAGACATGTGCTTACACACAAATATAGTACCATAGAAAAGTTGT
TGTCTAATTCAATTGTCAGACATGTGCTTACACACAAATATAGTACCATAGAAAAGTTGT
TTGTAATAATGTAACCTTGTAAATAGGTGCATAAACACTAATGCAATTGAAACAAAAGAAG
TGACATACACAATATAATCATATGTCTTCACACGTTGCTTACACGAGCAGCTCTGA
GGGTTCTGAAATCAATGTGGTCCCTCTTGCCCCTAAACAAAGATGGTTGTTGGGTTGG
ATTGACACTGGAGGCAGATAGTGCAGTAAAGTTAGTCTAAGGTTCCCTAGCTGTATTAGCCTCTG
ACTATATTAGTATAACAAAGAGGTATGTGGTTGAGACCAGGTGAATAGTCACATCAGTGTGGAG
ACAAGCACAGCACACAGACATTCTAGGAAGGAAAGGAACATGAAATCGTGTGAAATGGGTTGG
AACCCATCAGTGATCGCATATTCTCATGAGGGTTGCTTGAGATAGAAAATGGTGGCTCCTT
CTGCTTATCTCCTAGTTCTCAATGCTTACGCCCTGTTCTCAAGAGAAAGTTGTAACCT
CTGGTCTTCATATGTCCTGTGCTCCTTTAACCAAATAAAGAGTTCTGTTCTGGGGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSSLGAALLCGHGFRRVSGQKVCADFHKPCYKMFHELSSRVSFQEARNACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQTSGACPDLYQWSDGSNSQ
YRNWTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSKGRTKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-
145, 212-217